# High-Resolution, Unstructured Grids and Parallel Computing

**Mark Taylor**

**Sandia National Labs**

mataylo@sandia.gov

DCMIP Summer School, July 30-Aug 10, Boulder

U.S. Department of Energy
Office of Science

Sandia National Laboratories

# Outline

- Motivation for High Resolution
- Resolution and Parallel Computing
- Parallel Computing and unstructured grids
- Numerical methods for unstructured grids
- CESM parallel performance at high-resolution

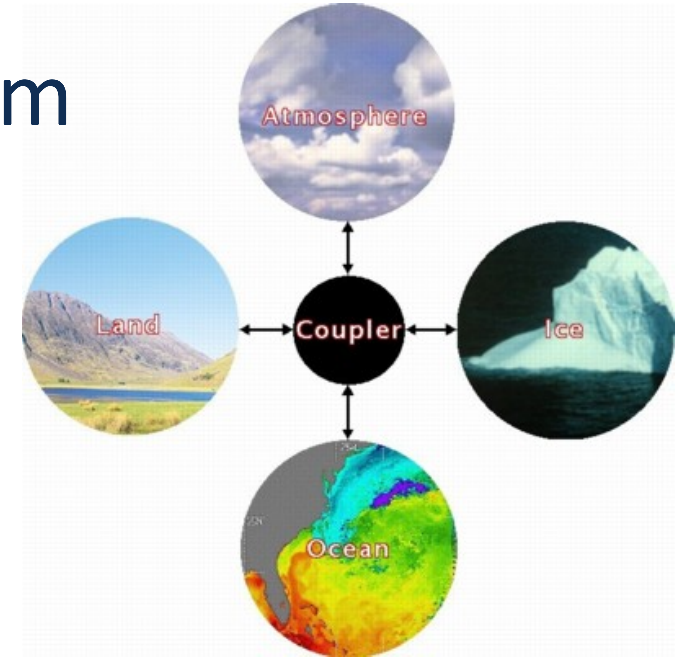
# Why High Resolution?

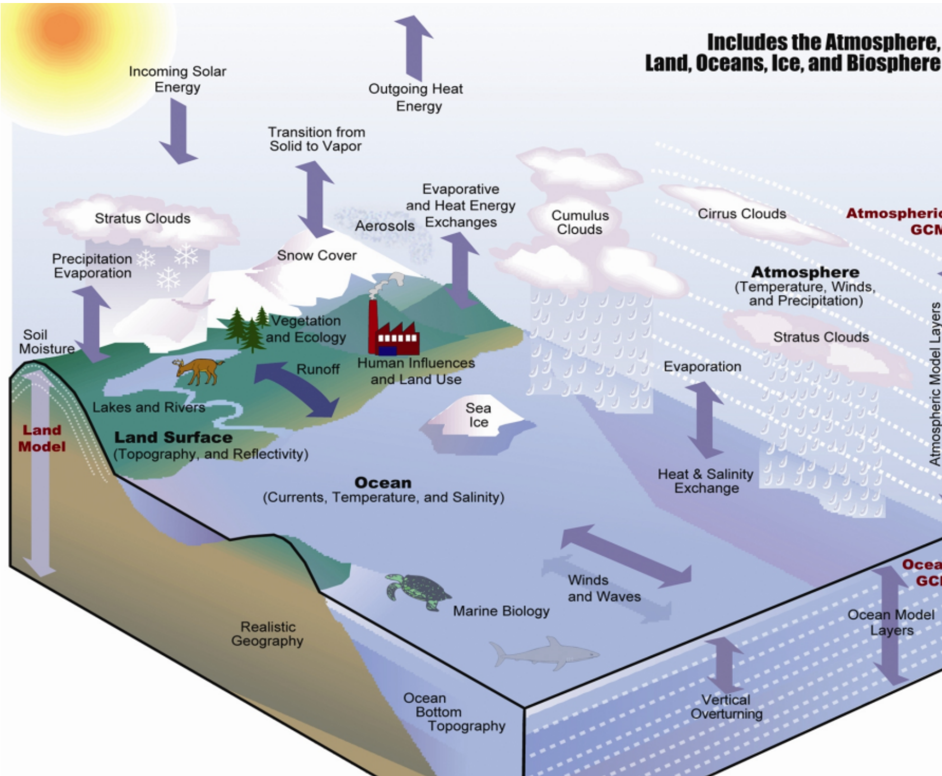
# Motivation for High-Resolution

- Run climate models at weather-forecast resolutions
- Regional climate simulations
- Resolve tropical cyclones in climate models
- Improved representation of climate variability including extreme events
- Resolve more processes, use less parameterizations
    - Cloud-resolving models: resolve moist convective processes, removing one of the largest sources of uncertainty in climate models

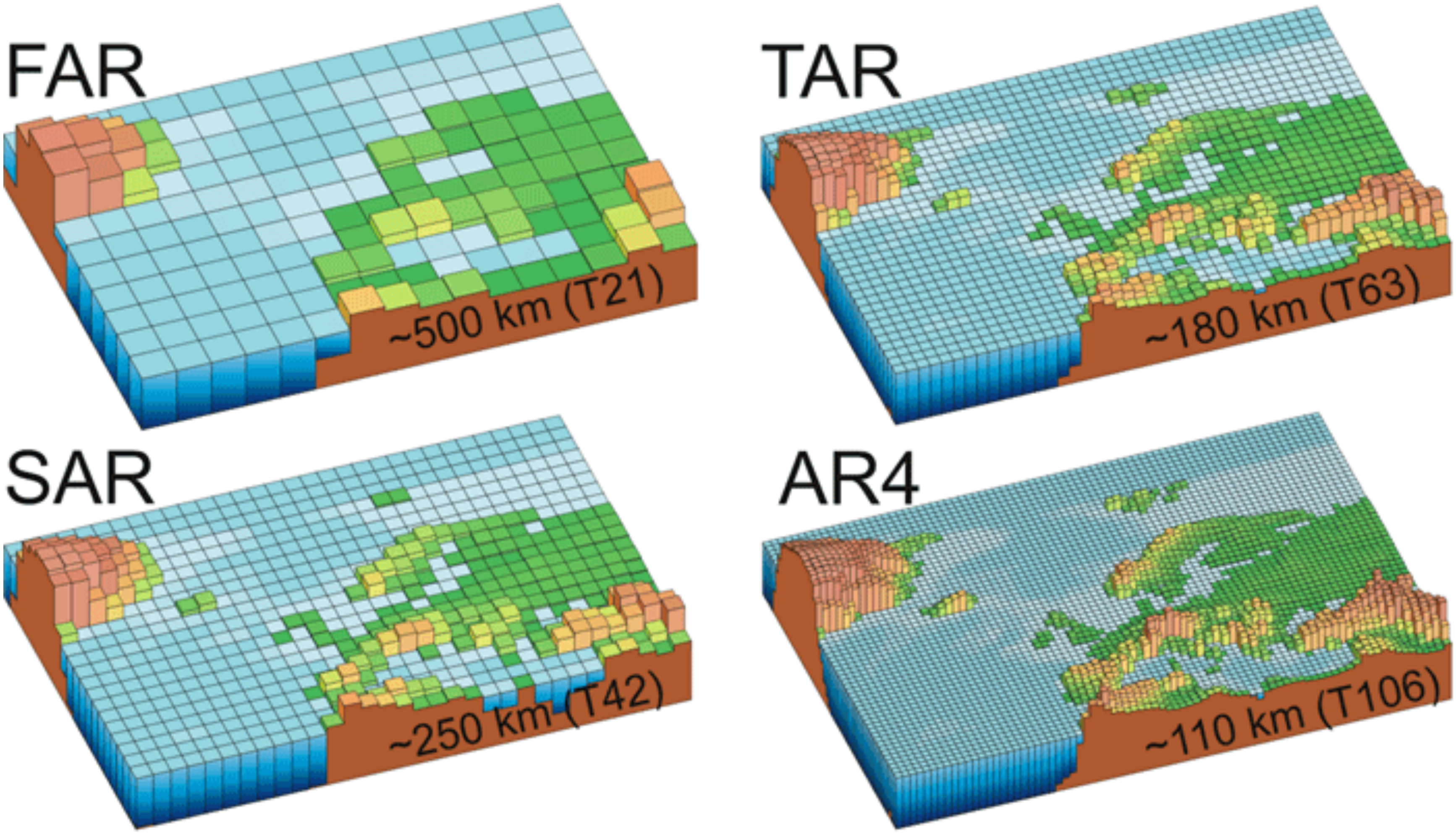
# The Community Earth System Model (CESM)

- CAM is the atmosphere component model for the CESM
- CESM is an IPCC-class model developed by NCAR, National Labs and Universities
- Atmosphere, Land, Ocean and Sea ice component models
- Science & policy applications:
  - Seasonal and interannual variability in the climate
  - Explore the history of Earth's climate
  - Estimate future of environment for policy formulation
  - Contribute to assessments



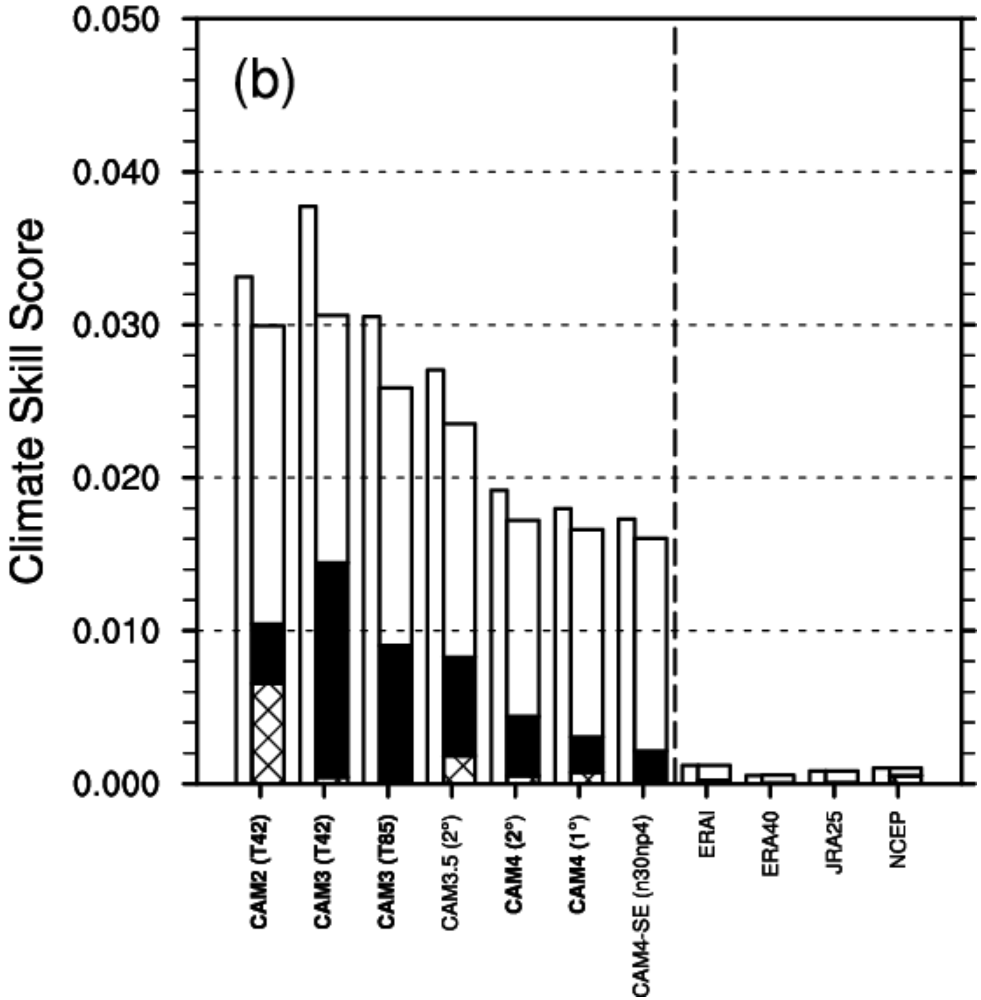



# Horizontal Grid Resolution



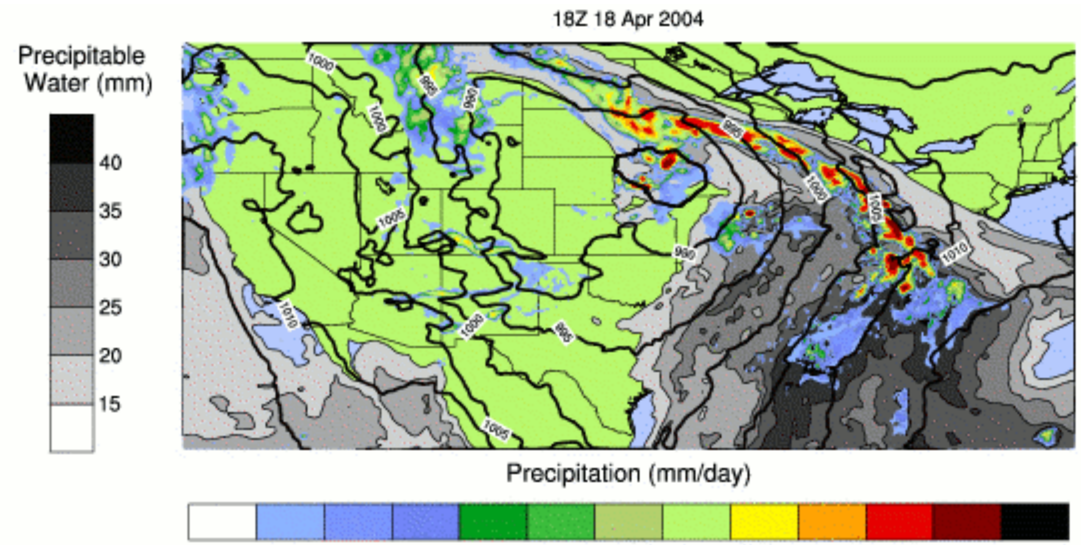

Source: IPPC 4th Assessment Report

# Improving Climate Skill Resolution and Physics



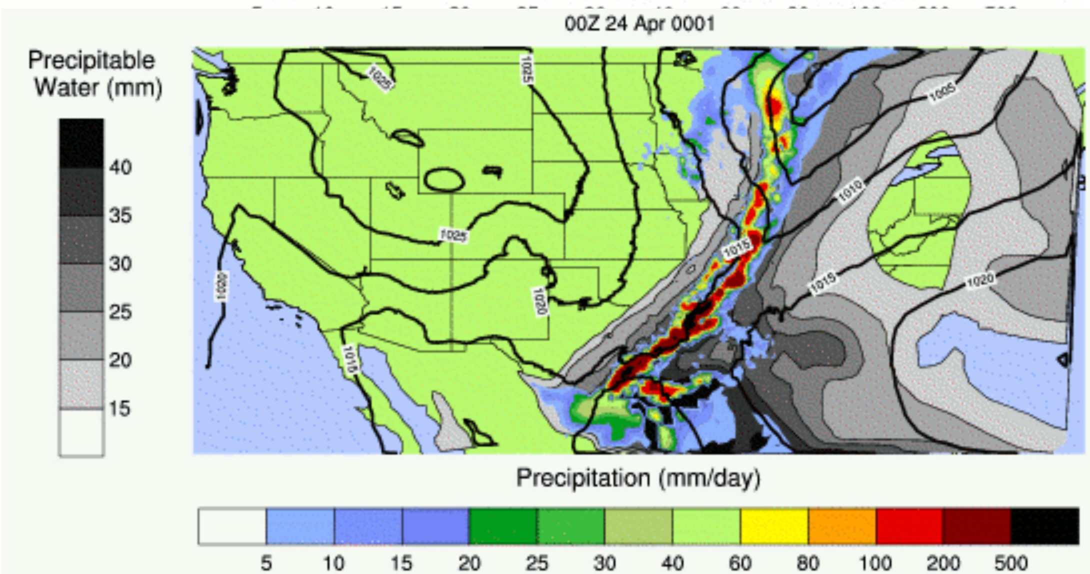

- CAM shows steady improvement as resolution is increased and physics improved
- 500mb geopotential height skill score (30-90N) DJF
- Mean square error from uncond. bias, cond. bias and phase error
- Source: Rich Neale (NCAR)

# Precipitable water (gray), precip rate (color), sea level pressure (contours)




## Global 1/8° Simulation

Snapshots show propagating convective system not seen at lower resolutions. Detailed frontal structure and tapping of moisture

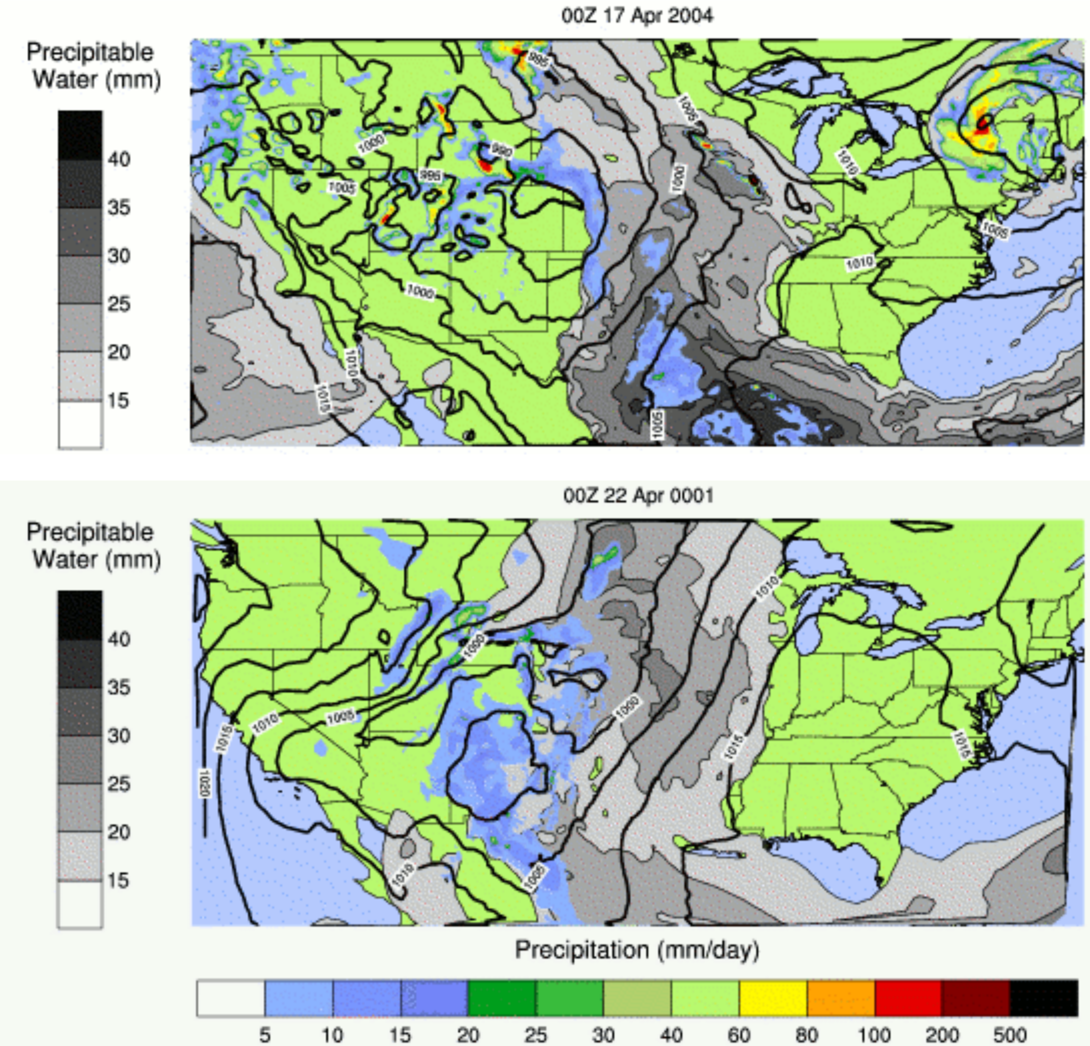


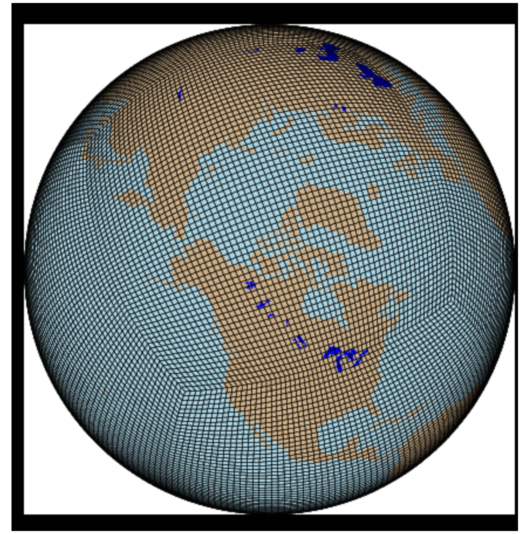

## Regionally Refined Simulation

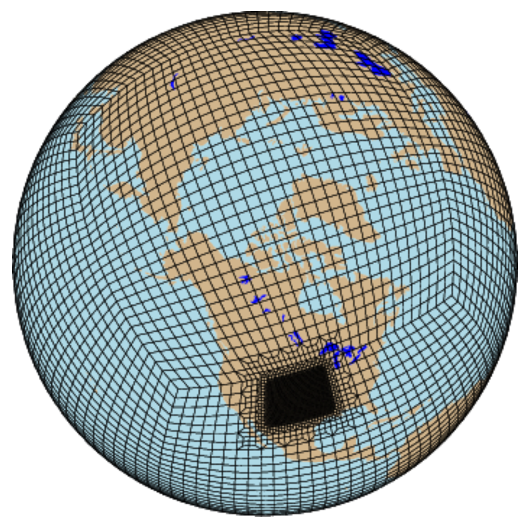
Similar convective systems form in the 1/8° region, strongly dissipated as it propagates into the 1° region

# Precipitable water (gray), precip rate (color), sea level pressure (contours)

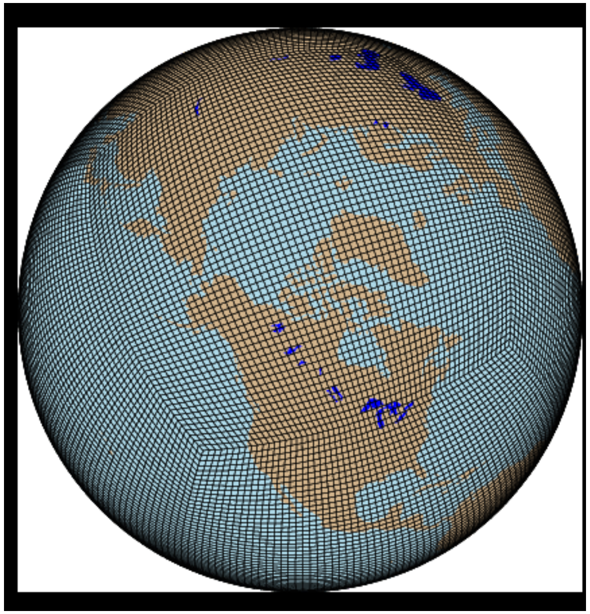
Sandia National Laboratories



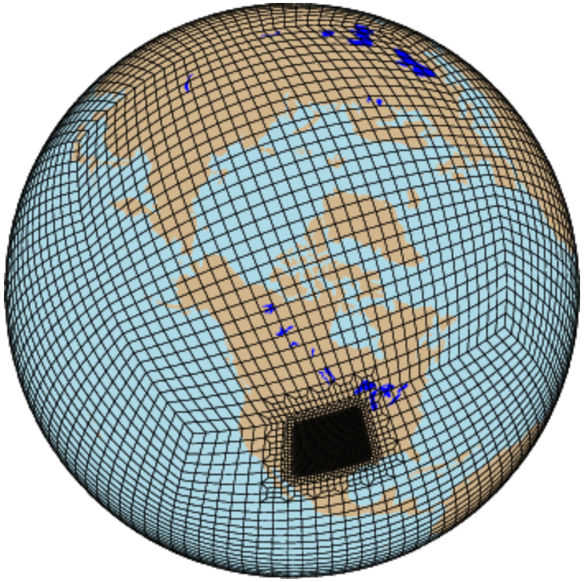


# CAM5-SE at 1/8°



**Global 1/8°**
CAM5-SE has a very efficient, scalable and *expensive* global 1/8° configuration.

- 6M core hours per year (ANL Intrepid)
- Yellowstone: 1-2M core hours?
- 3.1M physics columns
- dtime=600, dynamics dt=9.2

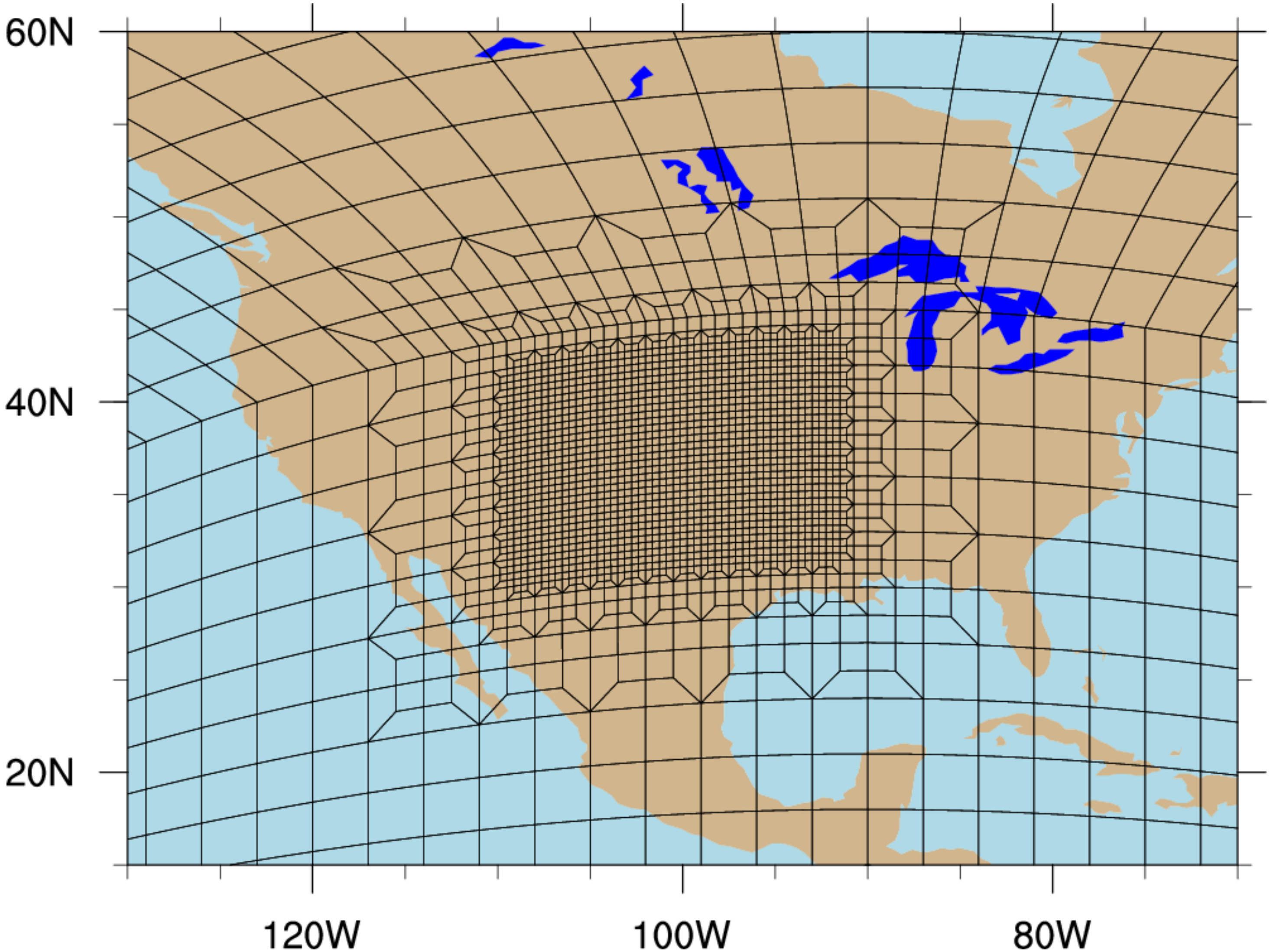


**SGP 8x Regionally Refined**
1° global resolution, refined to 1/8° continental sized region centered over SGP ARM site.

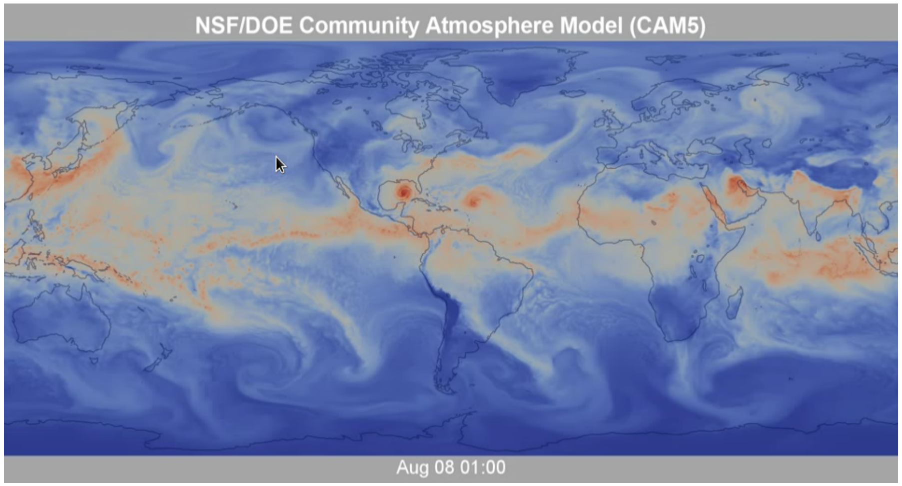
- 0.12 M core hours per year (Sandia Linux cluster).
- 67K columns.
- dtime=600, dynamics dt=7.9

# CAM5 Regionally Refined

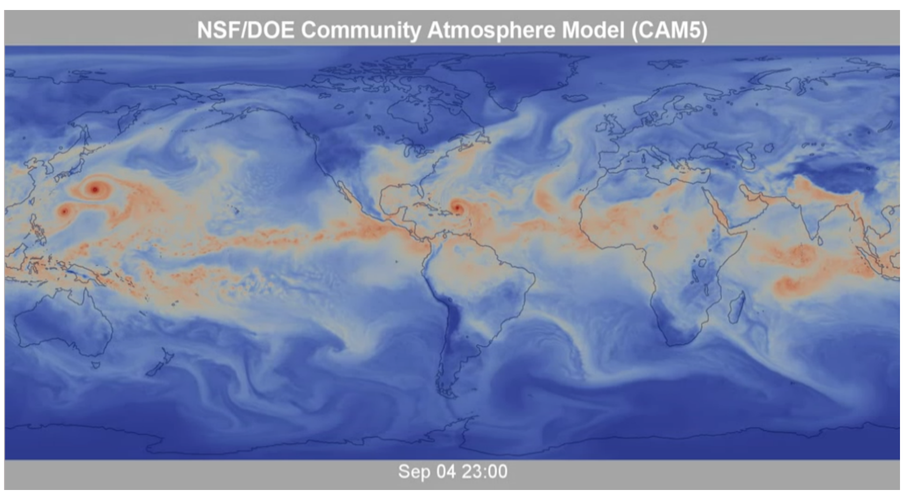
1° global resolution, refined to 1/8°



# CAM5-SE at 1/8°
# Precipitable water animations




**Category 5 storm in the Gulf of Mexico**

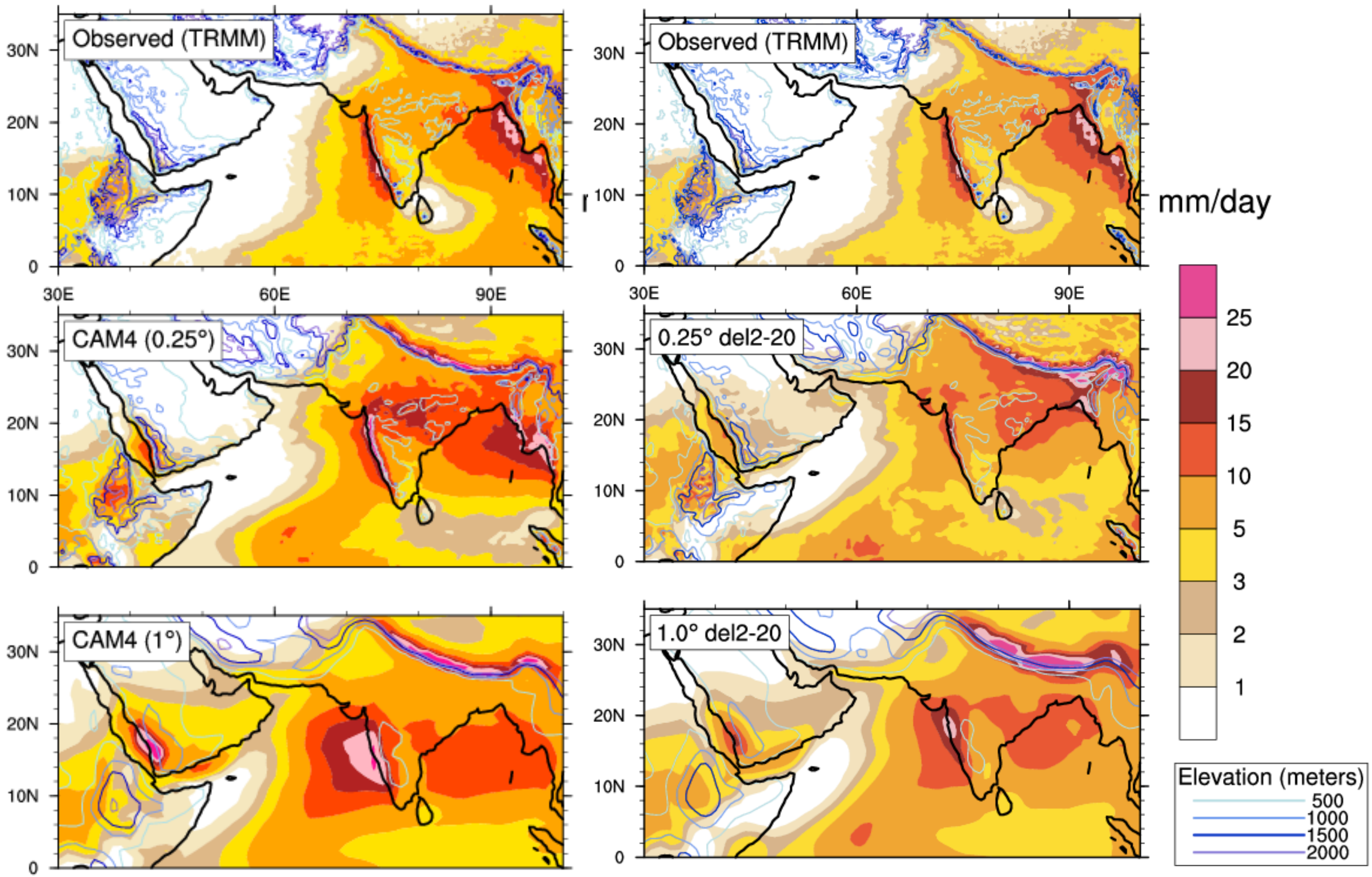



**Fujiwhara effect in the Pacific**

# Precipitation

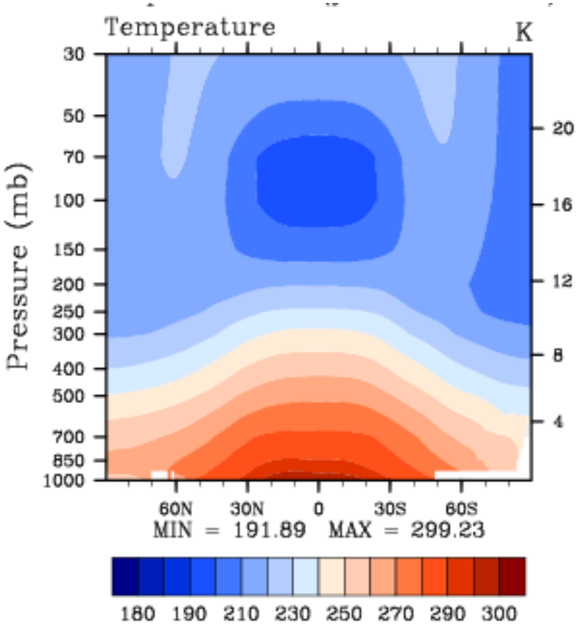
## CAM4-FV

## CAM5-SE

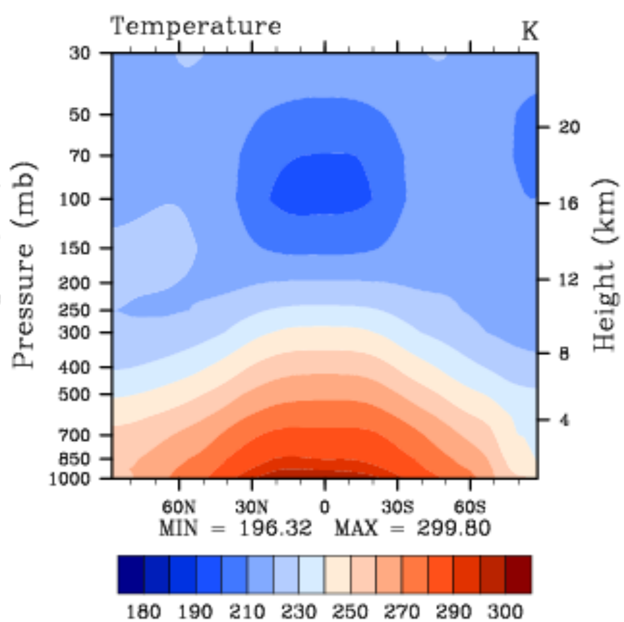


# Zonal Mean Temperature

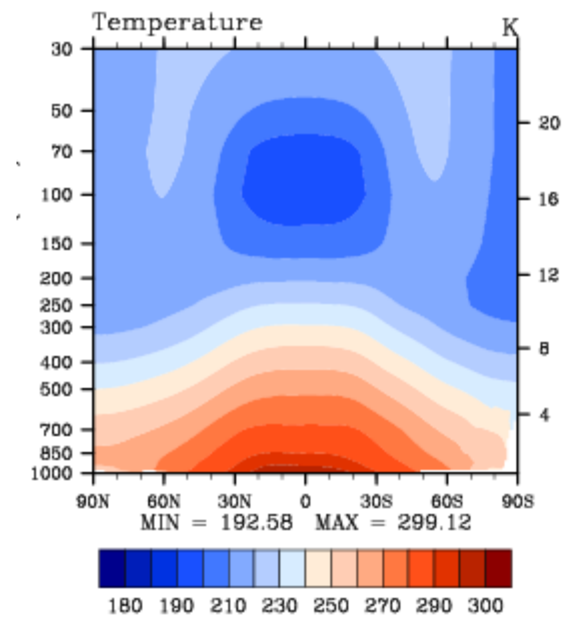
CAM4-SE 1°

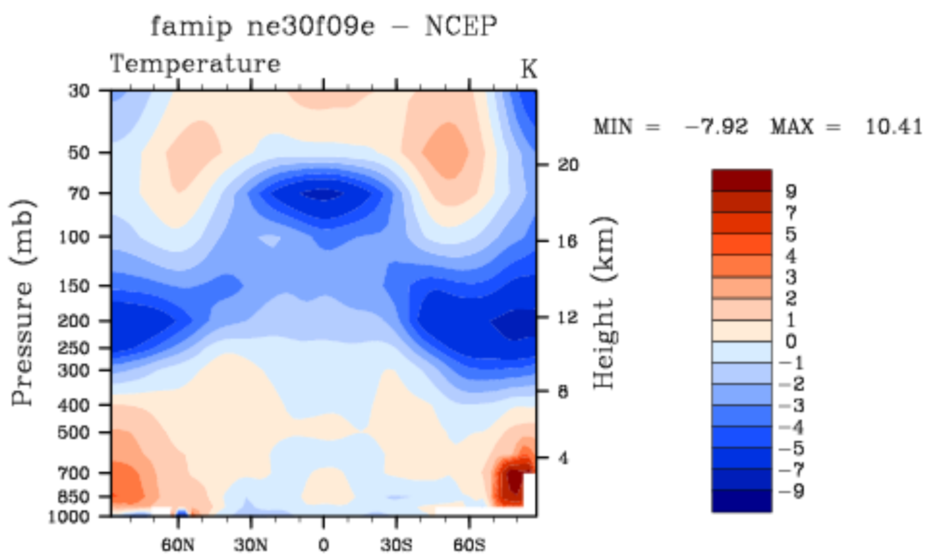



NCEP




CAM4-FV 1°

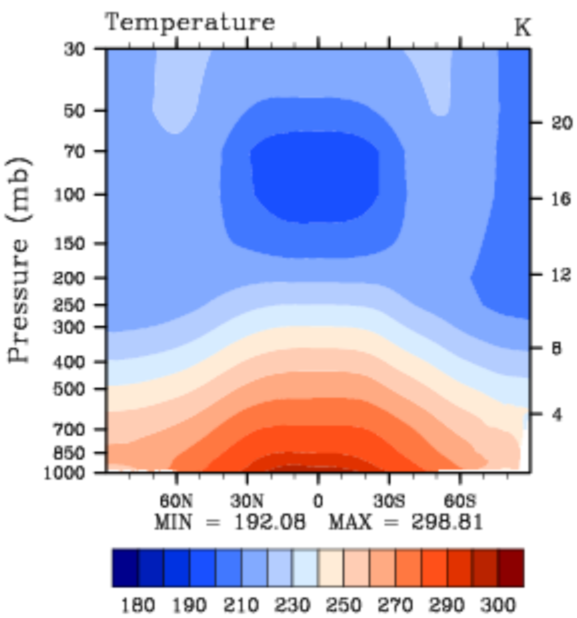


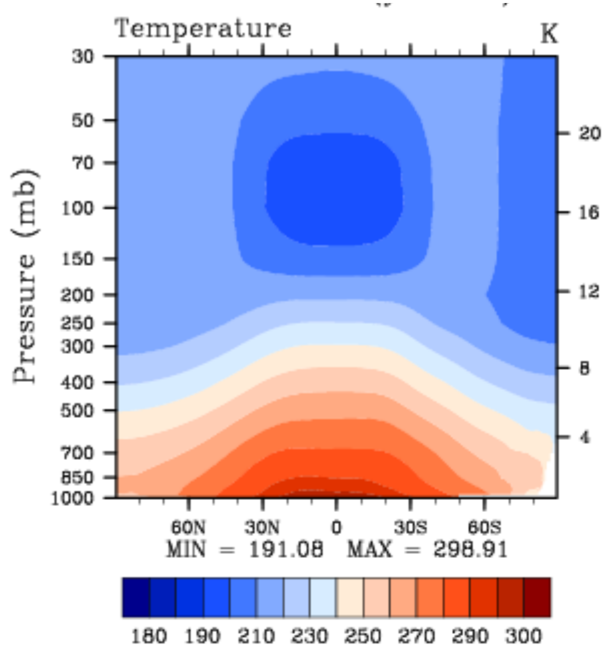

CAM has a long standing 200mb cold bias at the poles (both SE and FV)

# Zonal Mean Temperature

CAM4-SE 0.25°




CAM5-SE 1°

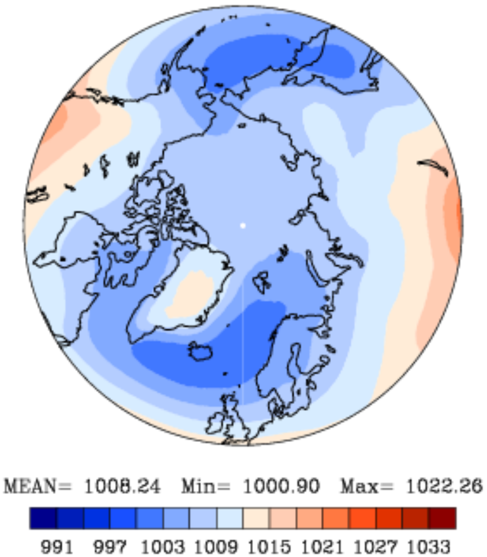



Result is insensitive to increasing resolution (left) or CAM5 physics (right)

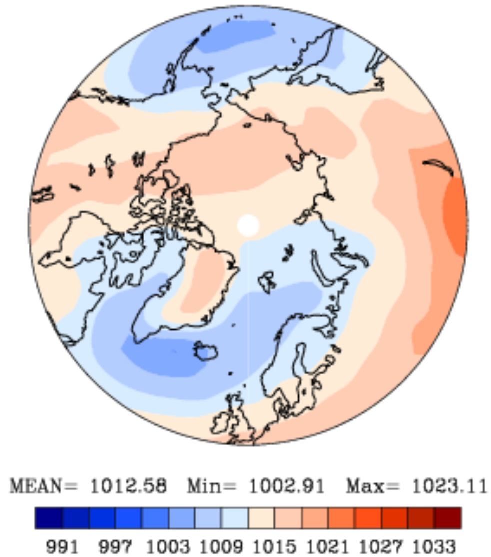
# Sea Level Pressure

Sandia National Laboratories

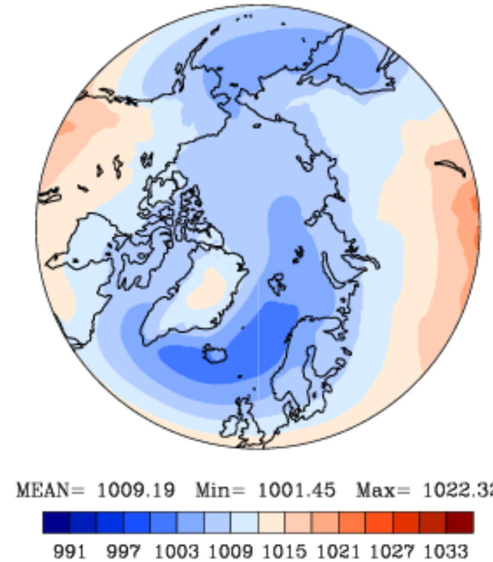
**CAM4-SE 1°**




**NCEP**




**CAM4-FV 1°**




**CAM4-SE 1/4°**

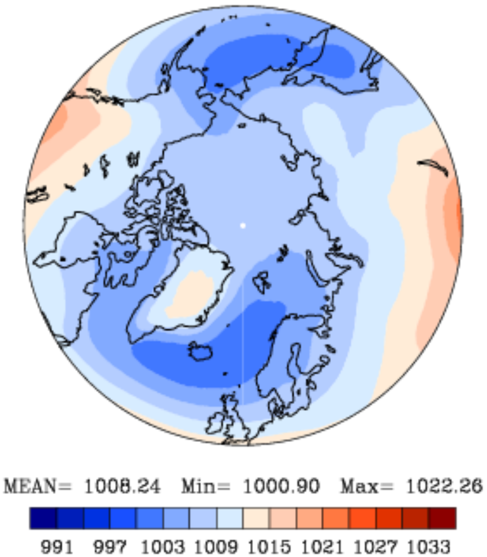
**CAM5-SE 1°**

CAM has too strong of an Icelandic low, in both SE and FV

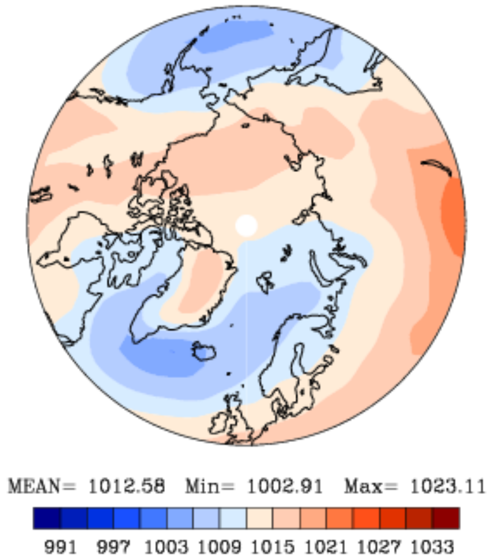
# Sea Level Pressure

Sandia National Laboratories

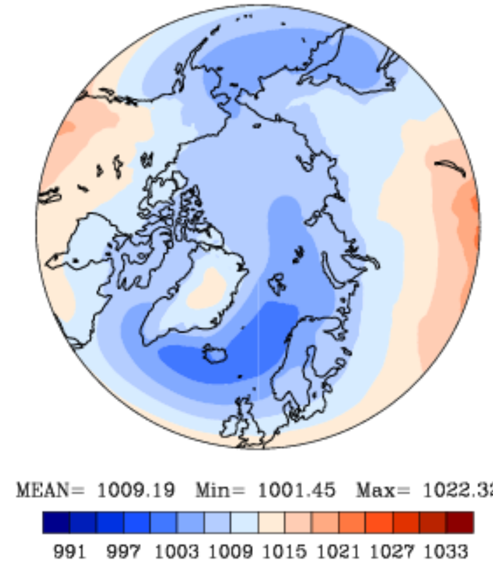
### CAM4-SE 1°



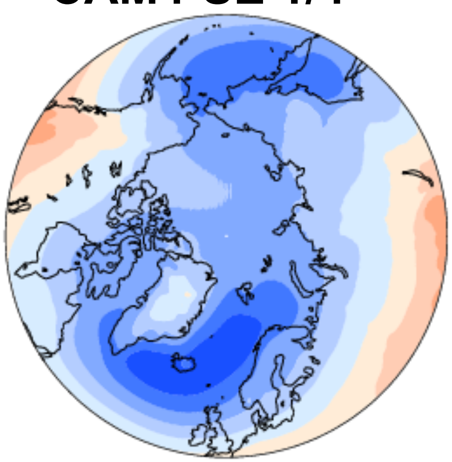

### NCEP



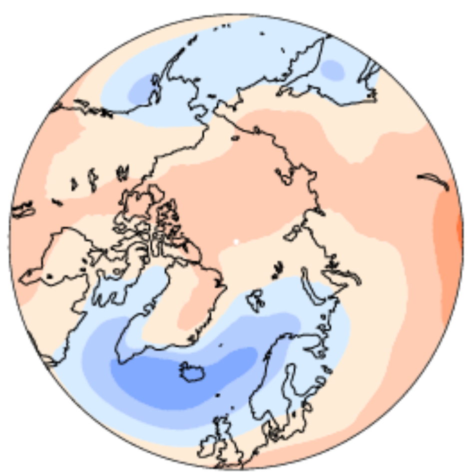

### CAM4-FV 1°




### CAM4-SE 1/4°



### CAM5-SE 1°

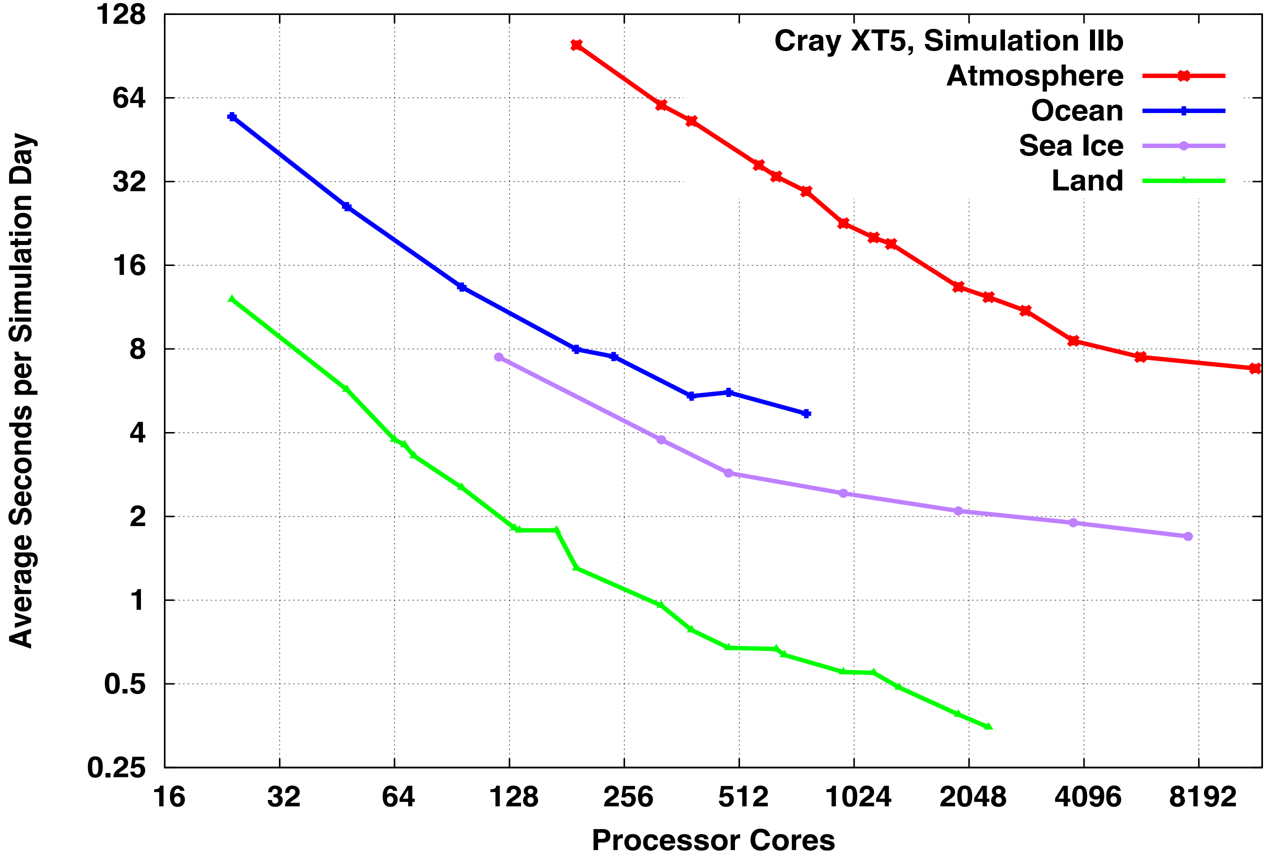


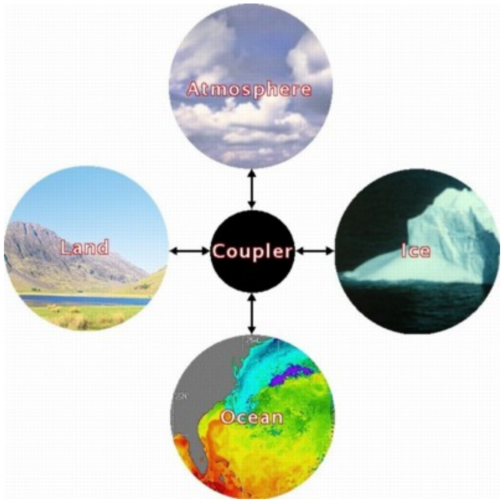
CAM has too strong of an Icelandic low, in both SE and FV

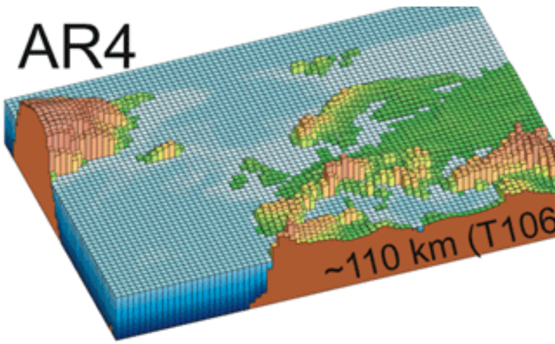
Icelandic low intensifies under mesh refinement, but is much improved with CAM5 physics

# Dynamical Core Scalability Bottleneck

# CESM performance (110km)

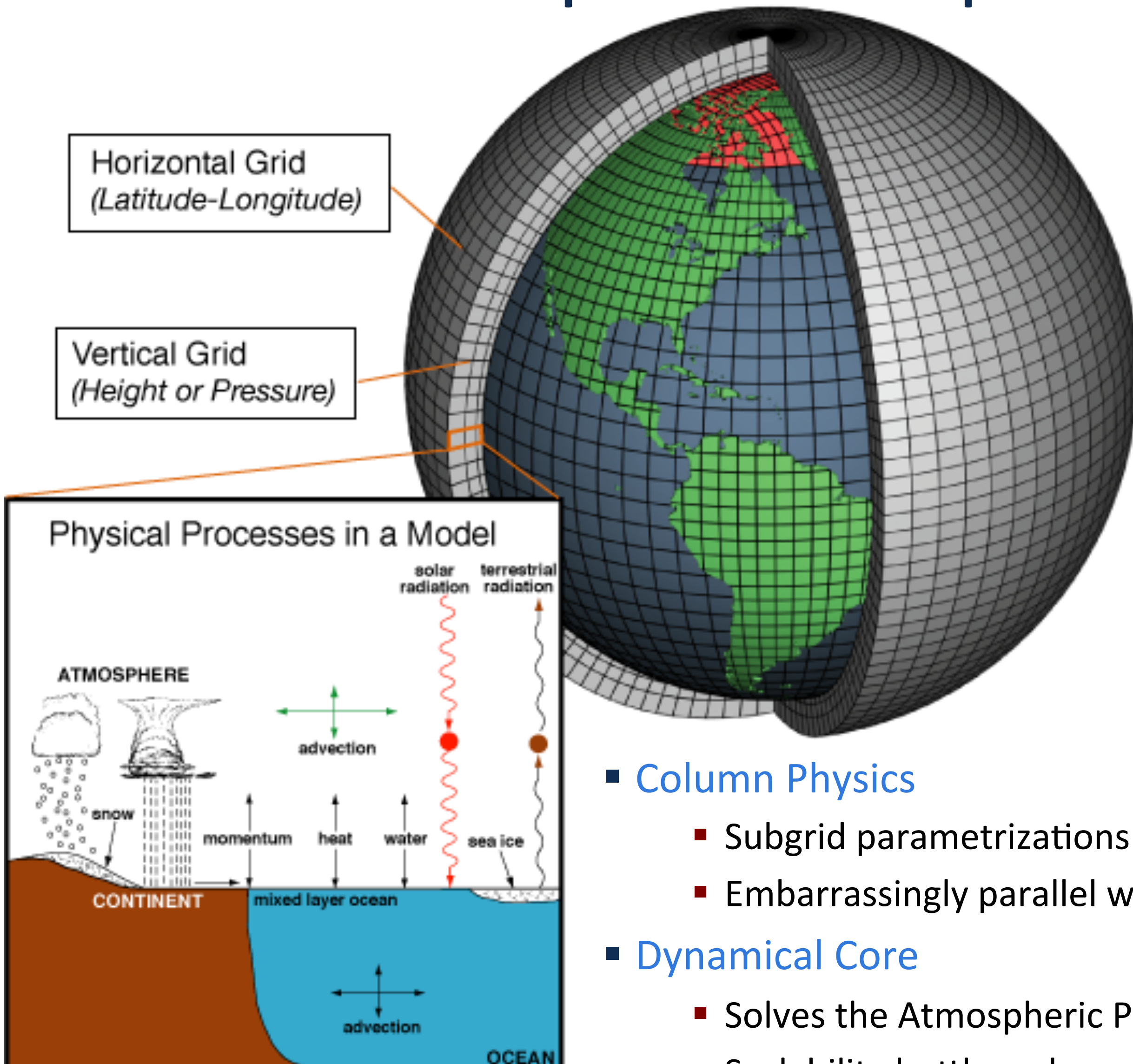



- CESM coupled simulation with all components at ~1° resolution
- Best observed component performance on the XT5
- Worley et al., Supercomputing 2011

# CESM Atmosphere Component (CAM)




- Column Physics
  - Subgrid parametrizations: precipitation, radiative forcing, etc.
  - Embarrassingly parallel with 2D domain decomposition
- Dynamical Core
  - Solves the Atmospheric Primitive Equations
  - Scalability bottleneck

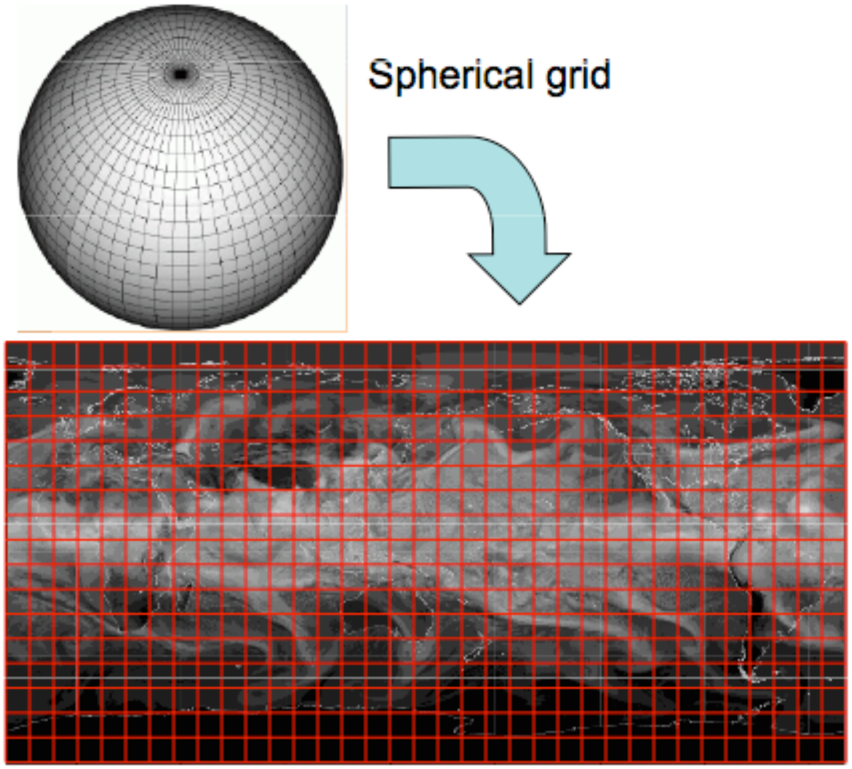
Source: http://celebrating200years.noaa.gov/breakthroughs/climate_model/welcome.html

# Dynamical Core Scalability Bottleneck



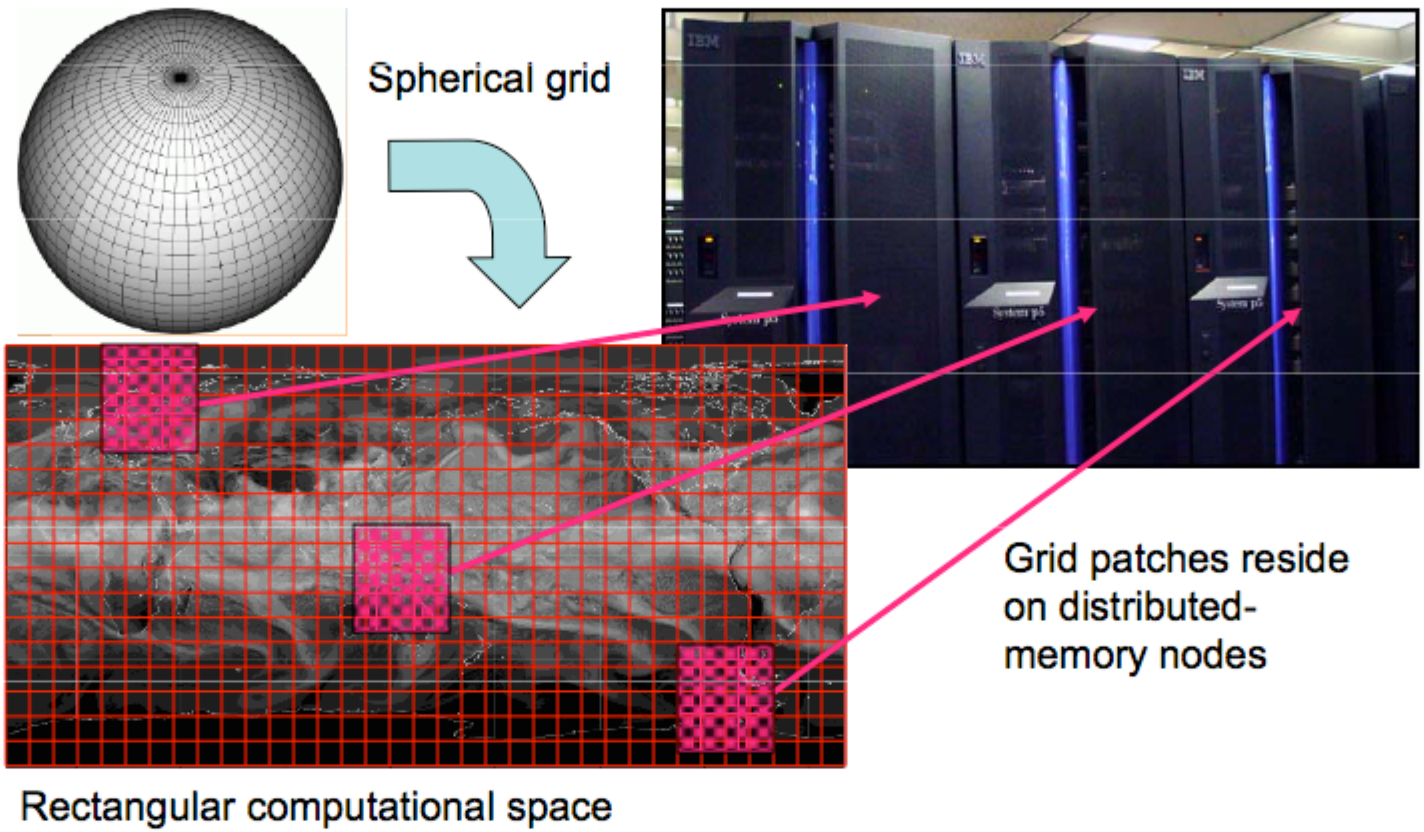
- Most dynamical cores in operational models use latitude-longitude grids:
  - Well proven.  Many good solutions to the "pole problem":  Spherical harmonics, polar filtering, implicit methods
  - But these approaches degrade parallel scalability

# The trouble with computers and our atmospheric cores…



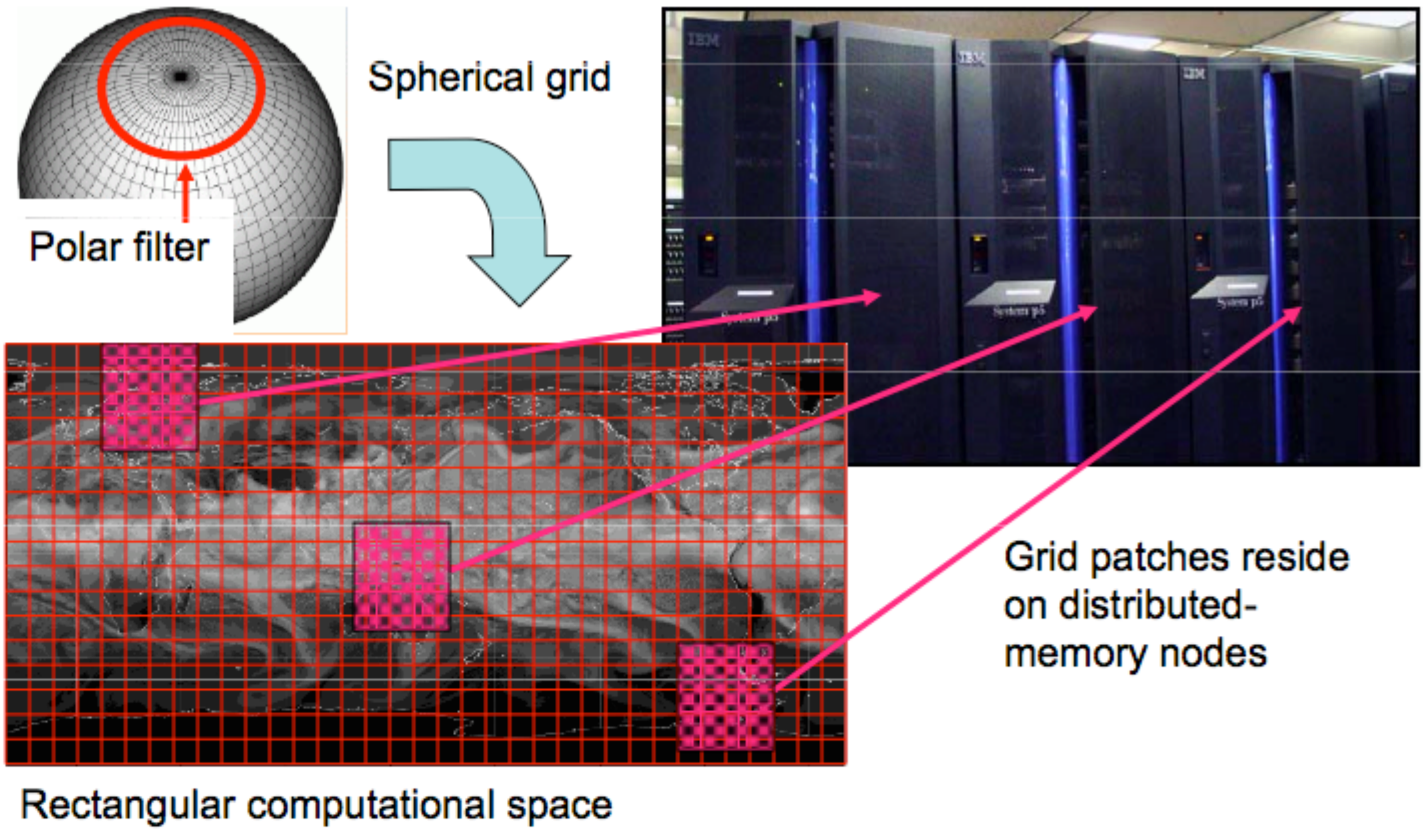

Source: Peter Lauritzen/NCAR

# The trouble with computers and our atmospheric cores…



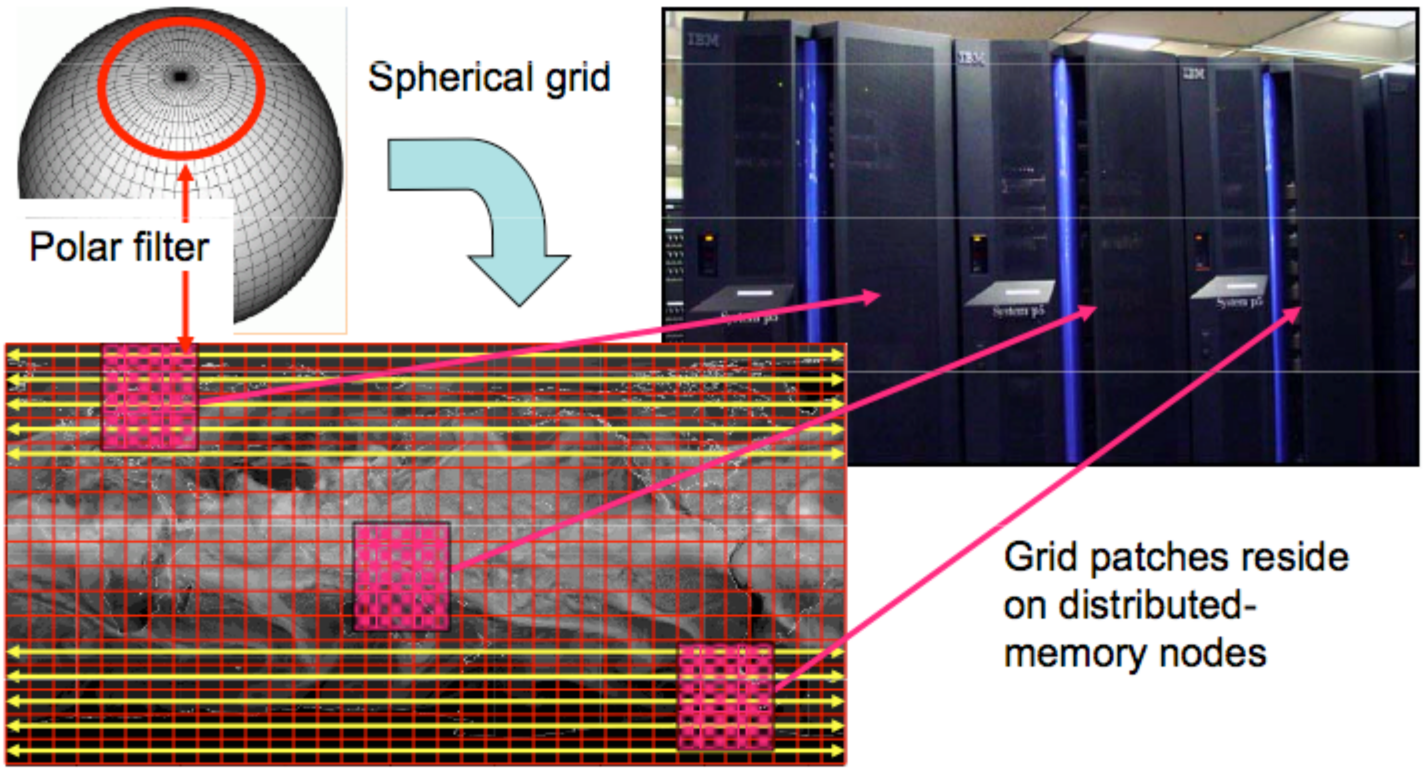

Source: Peter Lauritzen/NCAR

# The trouble with computers and our atmospheric cores…




Source: Peter Lauritzen/NCAR

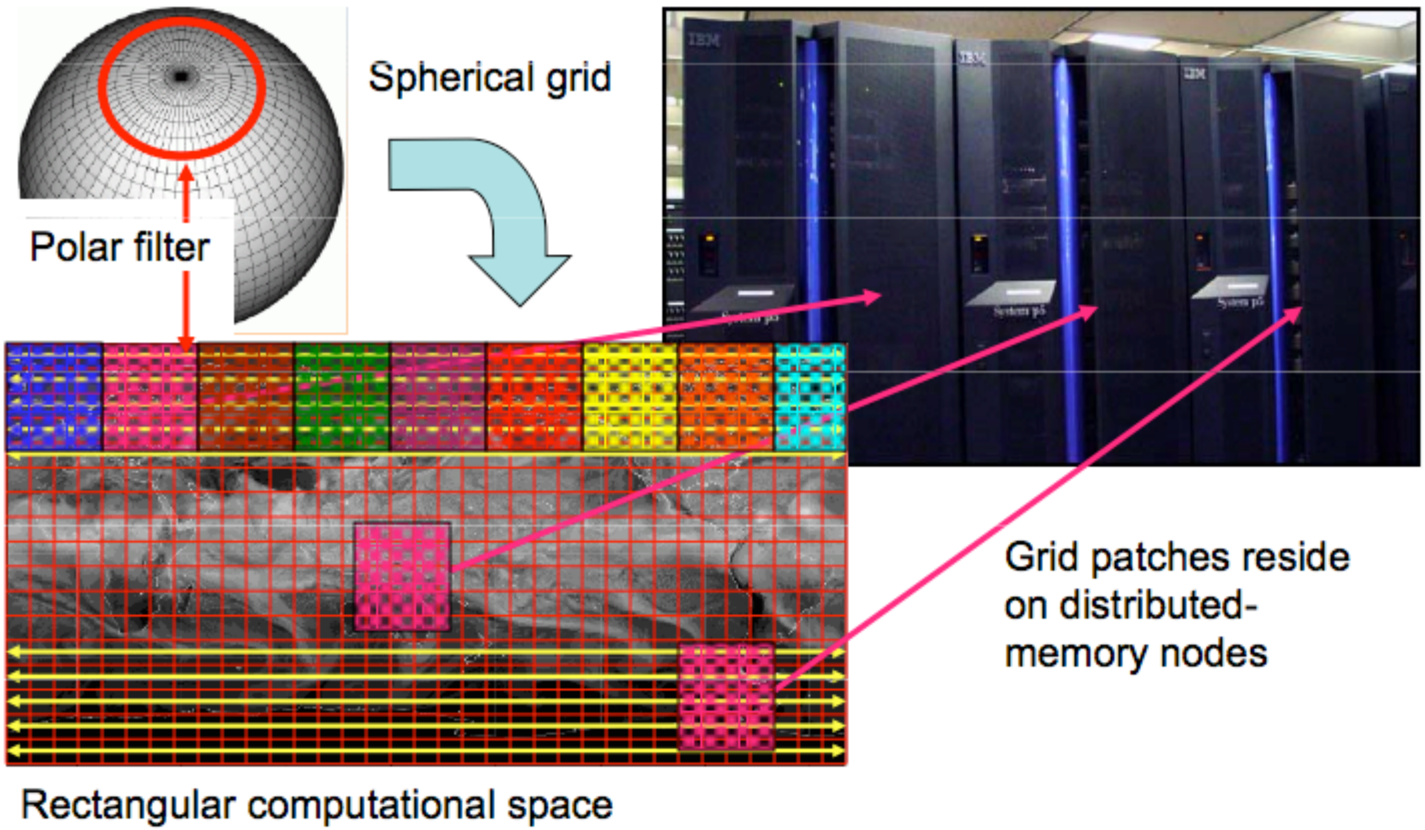
# The trouble with computers and our atmospheric cores…



Rectangular computational space

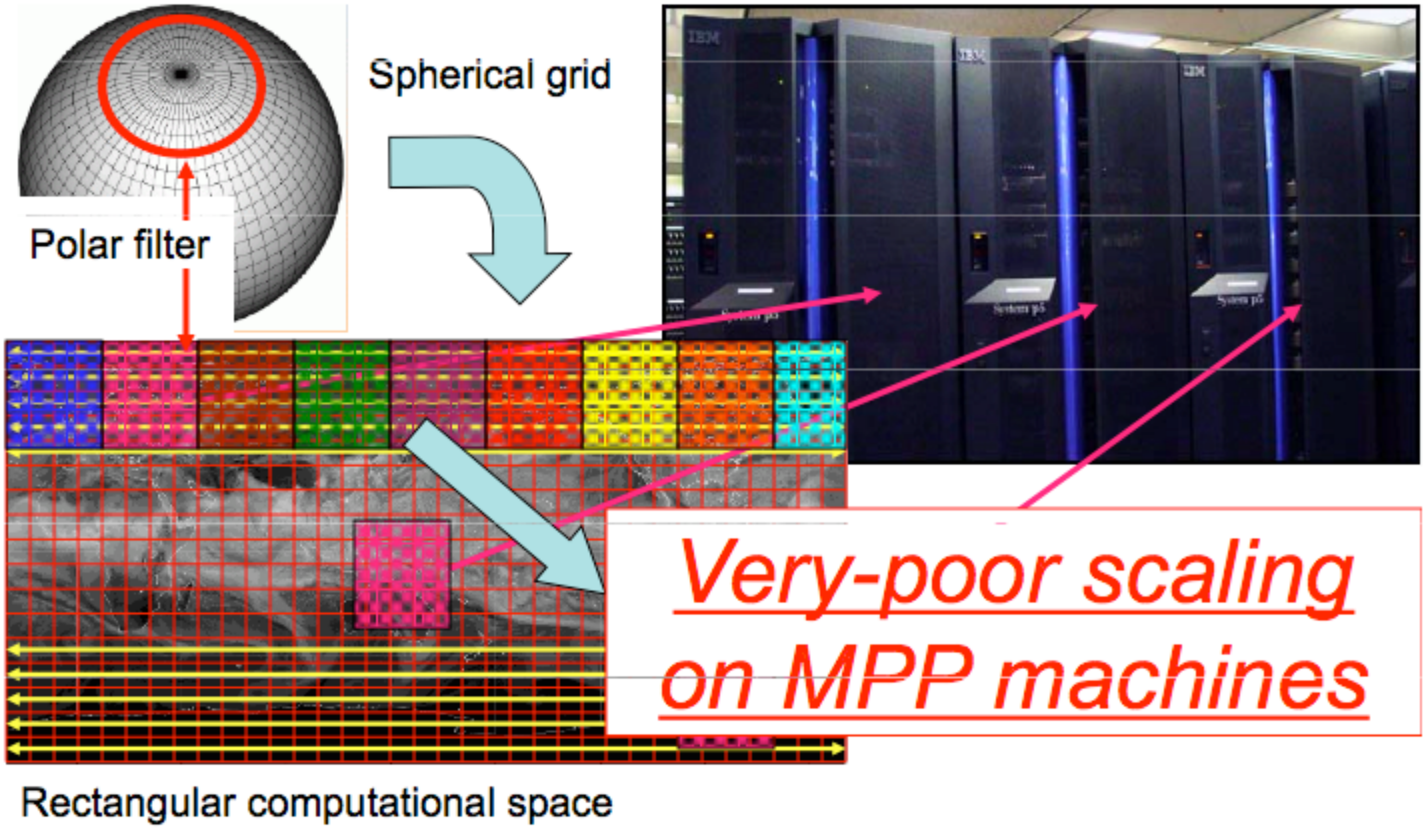
Source: Peter Lauritzen/NCAR

# The trouble with computers and our atmospheric cores…



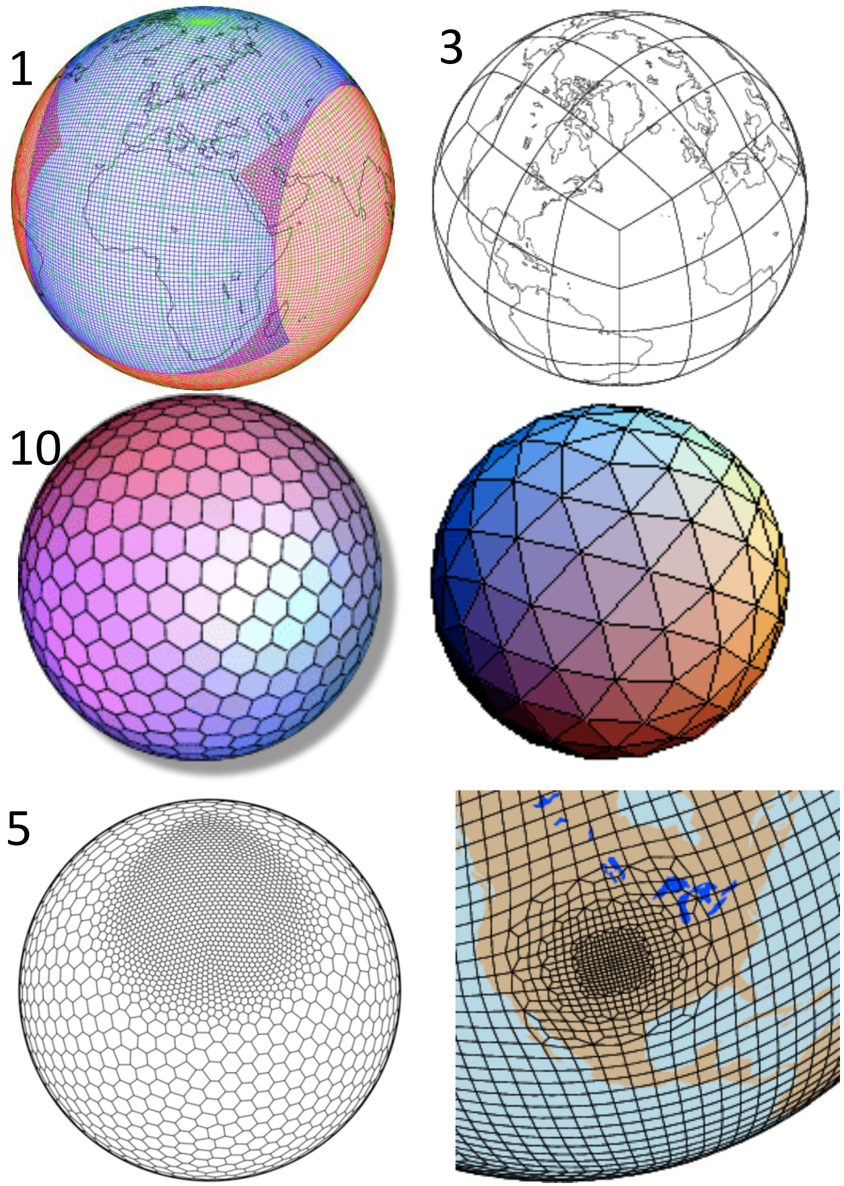

Source: Peter Lauritzen/NCAR

# The trouble with computers and our atmospheric cores…




Source: Peter Lauritzen/NCAR

# Atmospheric Dynamical Cores (Horizontal Grid)




- Quasi-uniform grids avoid the pole problem
- Can use full 2D domain decomposition in horizontal directions
- Each column in the vertical/radial direction kept on processor
- Equations can be solved explicitly, only nearest neighbor communication
- Requires more modern numerical methods
- Variable resolution and adaptive mesh refinement also possible

# Numerical Methods for unstructured grids

# Some History

- Brief history of the development of "block structured" dynamical cores
  - Follow one (of many) development threads based on block structured grids
  - Decompose the sphere into patches/elements of regions which are logically Cartesian
  - Early motivation: re-use methods developed for Cartesian grids and not have to deal with the pole problem
- Numerical issue: how do you patch together the different regions?

See also: Williamson, *The Evolution of Dynamical Cores for Global Atmospheric Models*, J. Meteor. Soc. Japan. 2007

# The Cubed-Sphere

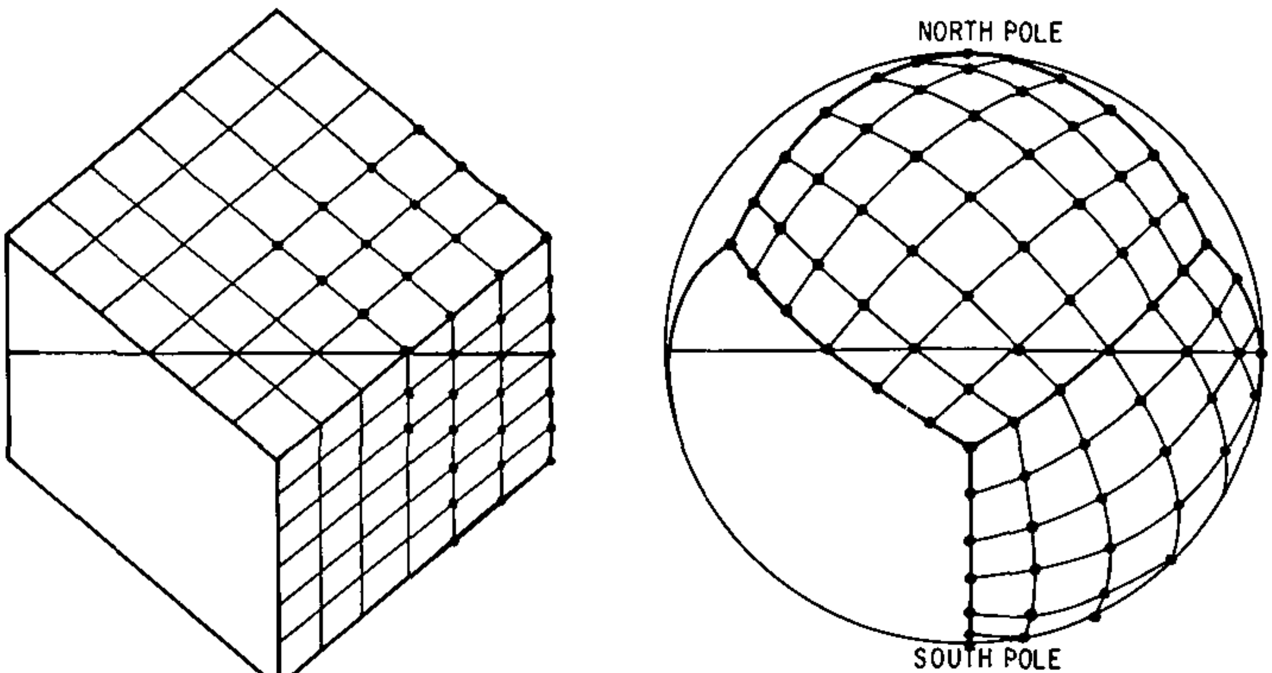



FIGURE 1.—A cubic representation of the earth. A cubic grid is shown together with the corresponding spherical grid which fits into the cubic splitting of the sphere, in the exact disposition that was used in the actual computations.

Source: Sadourny, MWR 1972

# Sadourny MWR 1972

- Used the Gnomonic projection (non-orthogonal)
- Finite difference method (mass & energy conserving)
- One sided differences used at cube face boundaries
- Large truncation error at the boundary resulting in noisy solutions
- Related approach: Phillips MWR 1959 (two stereographic polar caps + Mercator projection tropical band) with "missing" values for FD stencils obtained by interpolation

*Unfortunately, the decision to butt the coordinate systems together at a common latitude and to couple them with interpolation led to an unstable method, so the concept was abandoned.*
Browning, Hack, Swarztrauber MWR 1989.

# The Composite Mesh Method

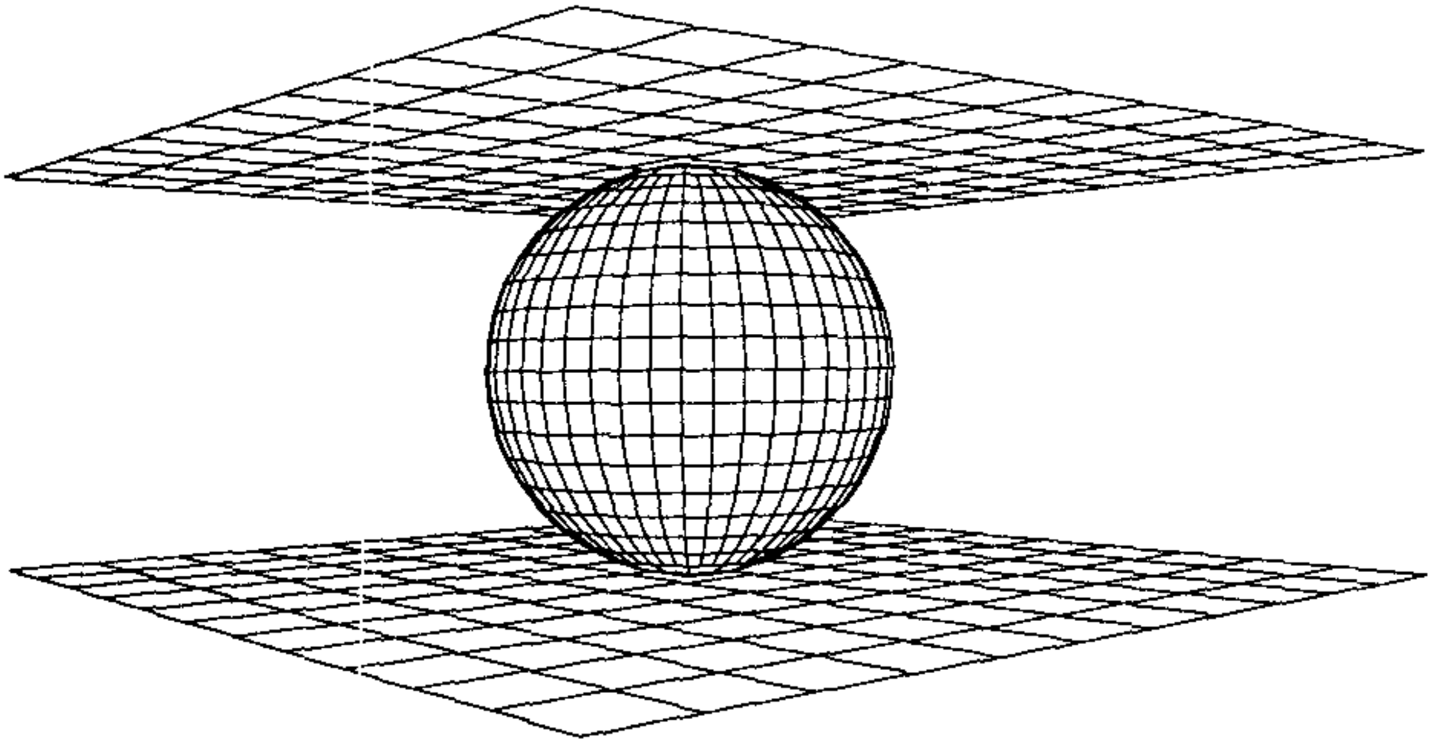


FIG. 1. The two tangent planes used in the composite-mesh method. Both planes extend slightly beyond the equator and contain a Cartesian coordinate system centered at the tangent point.

Source: Browning, Hack, Swarztrauber MWR 1989.

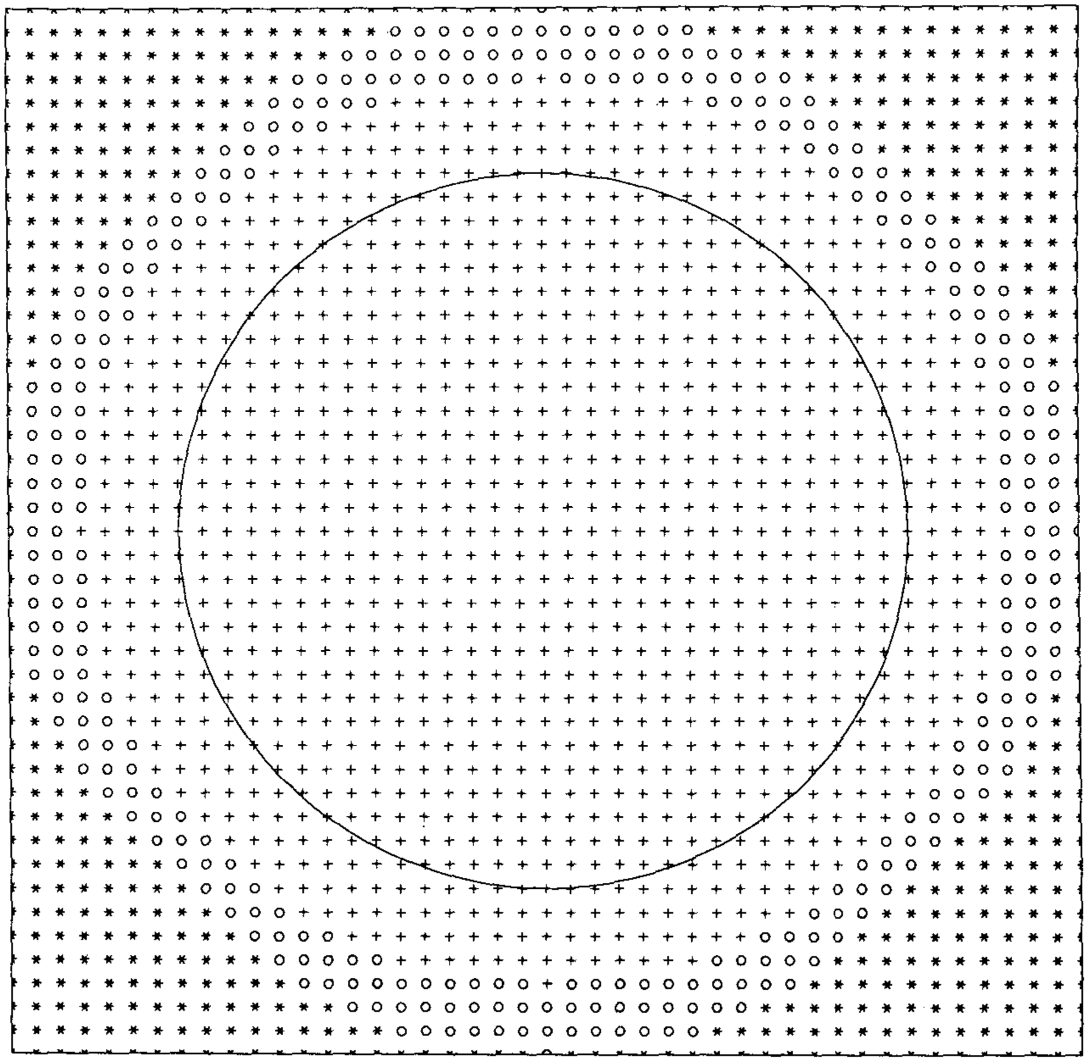


FIG. 2. The details of the composite grid mesh on either tangent plane for the case $N = 30$, $O = 7$, and $R = 3$. The circle is the image of the equator. The position of the points in $D$ and $I$ are marked by the symbols + and ○, respectively. The stars indicate points which are not used.

Browning, Hack, Swarztrauber MWR 1989.

# The Composite Mesh Method

- Stereographic (or other orthogonal) projections used so each patch maps to a regular Cartesian grid
- Boundary points from one grid (using one coordinate system) are interior points from another grid (using a different coordinate system)
- The overlapping of all boundary points is the key to the stability of the method (Starius, Numer. Math. 1977,1980)

*...there is an overlapping of the grids in the middle latitudes, and one needs to interpolate values from one grid to its neighbor in the course of the calculation. This need makes the design of a global conservative scheme impossible in practice.*

Sadourny, MWR 1972

# The Composite Mesh Method

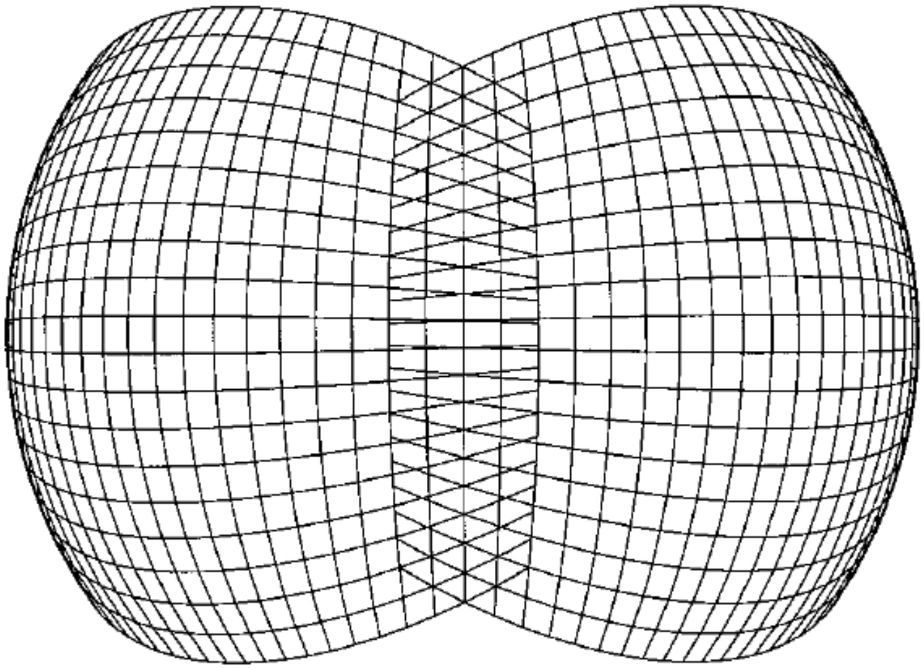


**FIG. 4.** View of two contiguous equatorial blocks and two of their stencils. The view is centered on the common vertical boundary line and shows the case of two grids with $N_\eta = N_\xi$ and $N_S = 2$. Notice that in this case, since both blocks use the same grid spacing, the horizontal coordinates of the stencil points of one grid coincide with the horizontal coordinates of the last two interior grid points of the contiguous block.

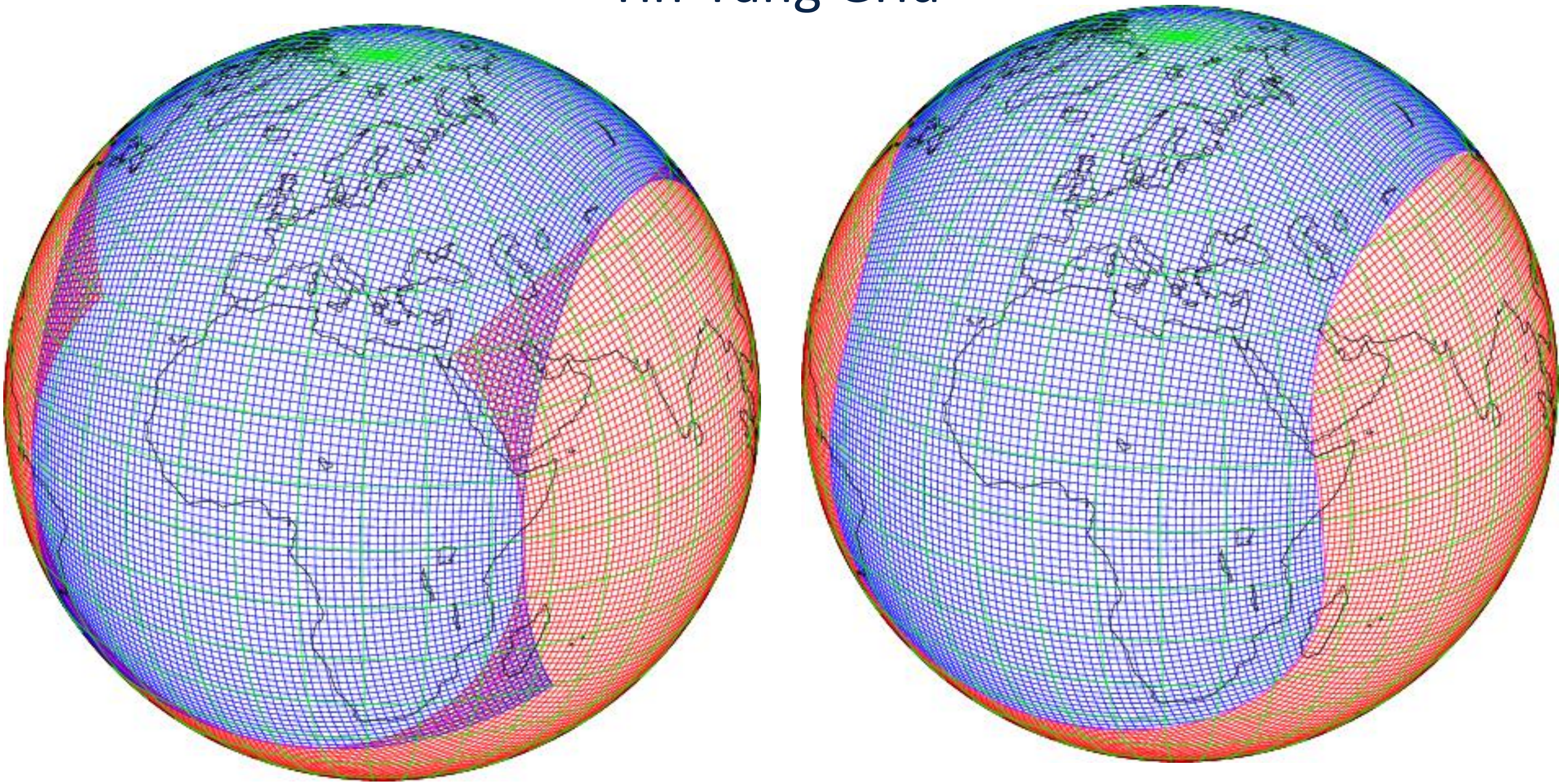
Ronchi, Iacono, Paulucci, JCP 1996

# The Composite Mesh Method

- Ronchi, Iacono, Paulucci JCP 1996
- First use of the phrase “cubed-sphere”?
- 4th order, fully co-located A-grid like method

# The Composite Mesh Method
# Yin-Yang Grid

Sandia National Laboratories



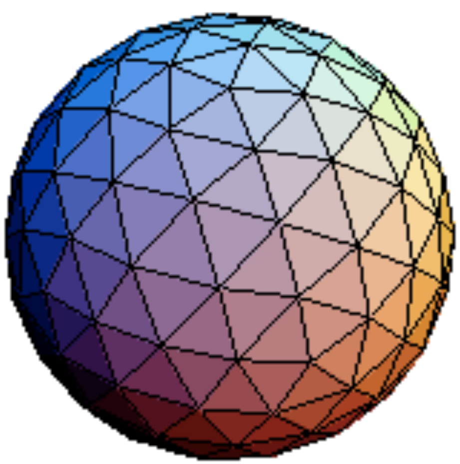
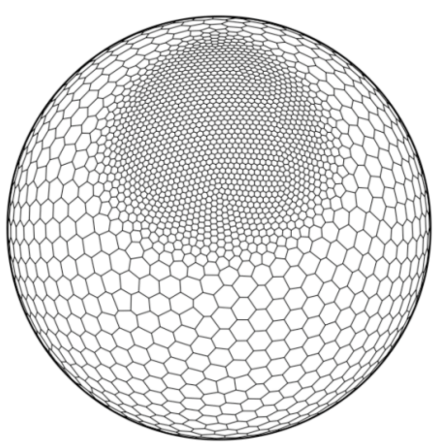
Source: R.J. Purser (NCEP) *The bi-Mercator Grid as a Global Framework for Numerical Weather Prediction*

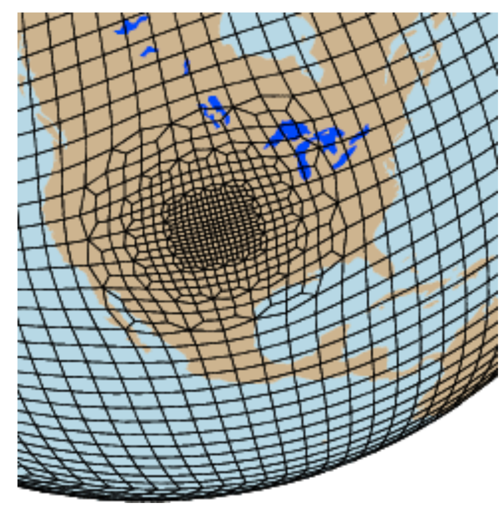
Kageyama, Sato, Geochem. Geophs. Geosyst. 2003

# Non-overlapping grids

# Methods for Unstructured Grids

- Modern FV and FE methods designed for fully unstructured grids
  - Used for grids with some structure (icosahedral, cubed-sphere)
  - In principal work on fully unstructured grids
- Finite Volume methods well represented at DCMIP 2012
- Finite Element methods are less common (CAM-SE)



# Finite Elements

- FE represent another approach to solving Sadourny's grid coupling problem
  - Derivatives and other terms are computed locally using only element data, independent of their neighbors
  - Elements are coupled together using a Galerkin approach
- Galerkin approach leads to an implicit system – inverting the FE *mass-matrix*, making them expensive for time-dependent problems
- Two popular approaches which lead to a diagonal mass matrix: Discontinuous Galerkin and Spectral Finite Elements

# Galerkin FE Methods
# Example: Tracer Advection

Advection Equation:

$$\frac{\partial q}{\partial t} + \nabla \cdot (q\vec{u}) = 0$$

Forward Euler:
(or convex combinations like SSP RK)

$$q(t+1) = q(t) - \Delta t \nabla \cdot (q(t)\vec{u}(t))$$

Multiply by test function $\phi$ and integrate over the sphere $\Omega$:

$$\int_\Omega \phi q(t+1) = \int_\Omega \phi q(t) - \Delta t \int_\Omega \phi \nabla \cdot (q\vec{u})$$

# Galerkin FE Methods
# Example: Tracer Advection

Decompose over elements (still exact)

$$\sum_m \int_{\Omega_m} \phi q(t+1) = \sum_m \int_{\Omega_m} \phi \left( q(t) - \Delta t \, \nabla \cdot (q\vec{u}) \right)$$

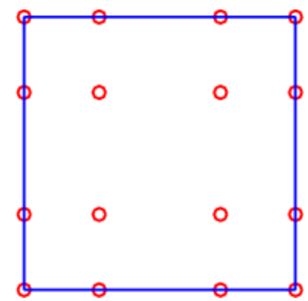
Given a velocity field and a solution *q* to the advection equation, this equation must hold for any reasonable function $\phi$

Galerkin discretization:  instead of asking that the original equation hold at N grid points, we instead solve this system of integral equations for N test functions $\phi$

# Galerkin FE Methods
# Example: Tracer Advection

Descretization: replace integrals by quadrature. (For quadrilateral elements, Gauss-Lobatto quadrature is very efficient)

$$\sum_{m} \sum_{\Omega_m} \phi q(t+1) = \sum_{m} \sum_{\Omega_m} \phi \left( q(t) - \Delta t \, \nabla \cdot (q\vec{u}) \right)$$



Let $H_1$ = space of globally continuous functions which are polynomials of degree p in each element

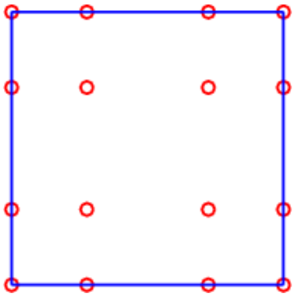
Let $\phi_n$ be a basis for $H_1$

Find q in $H_1$ which solves the above equation for all $\phi_n$. Note: divergence can be computed "exactly"

# Galerkin FE Methods
# Example: Tracer Advection

Solution Procedure:

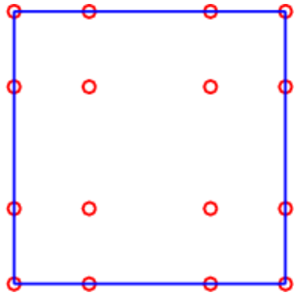
Step1: compute within each element:



$$q^* = q(t) - \Delta t \nabla \cdot (q\vec{u})$$

Step 2: Find q in $H_1$ which solves, for all $\phi_k$

$$\sum_m \sum_{\Omega_m} \phi_k q(t+1) = \sum_m \sum_{\Omega_m} \phi_k q^*$$

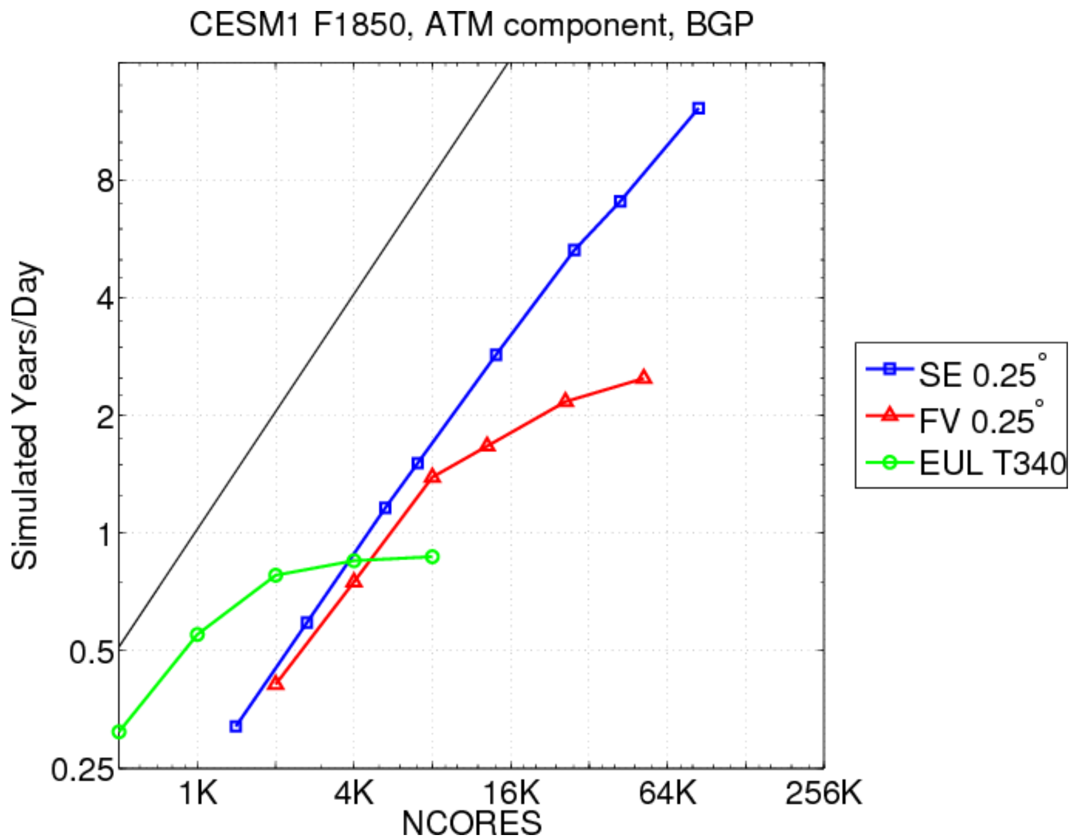


# Galerkin FE Approach Ideal for Modern Architectures

- Galerkin formulation of the equations leads to a 2 step solution procedure:
  - Step 1: All computations local to each element and on a tensor-product grid. Structured data with simple access patterns and arithmetically intensive operations: Extremely efficient on modern CPUs or GPUs
  - Step 2: Apply inverse mass matrix (projection operator).
- All inter-element communication is embedded in Step 2, providing a clean decoupling of computation & communication.
  - Only a single routine has to be optimized for parallel computation.
  - Gordon Bell Awards: 2000 (best performance, NEK5000), 2001 (honorable mention, HOMME) , 2003 (best performance, SPECFEM3D)

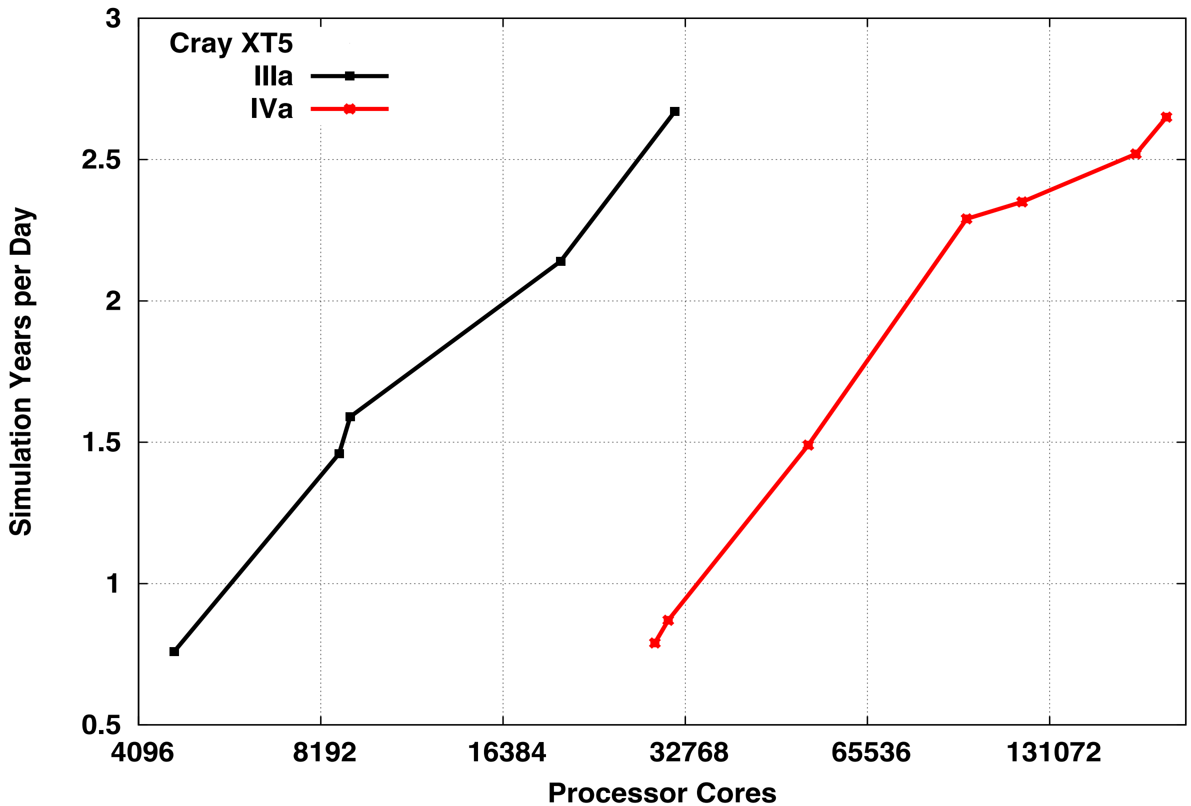
# CAM4
# Parallel performance and scalability

# CAM4 0.25° (28km) Scalability
# IBM BG/P "Intrepid"




- Compare CAM with SE, FV and EUL (global spectral) dycores
- CAM-SE achieves near perfect scalability to 1 element per core (86,000 cores). Peak performance: 12.2 SYPD.
- Atmosphere only times. Full CESM runs ~50% slower because of other components

# CESM Fully coupled configurations
# Cray XT5 "Jaguar"



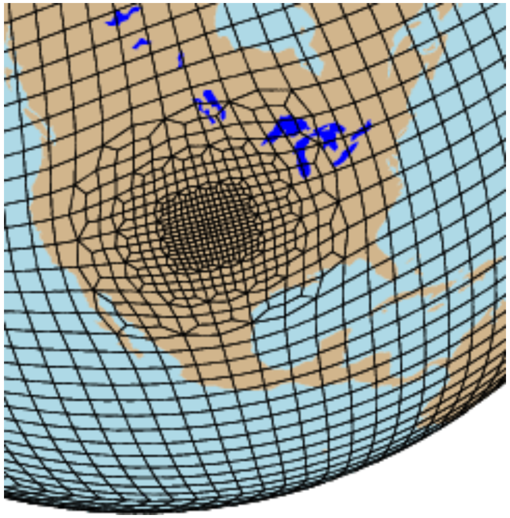

- Black: 0.25° (28km, CAM4-FV) Atmosphere, 0.1 (10km) Ocean:  Running at 2.6 SYPD on 32,000 cores
- Red: 0.125° (14km, CAM4-SE) Atmosphere, 0.1 (10km) Ocean:  Running at 2.6 SYPD on ~150,000 cores.
- Worley et al., Supercomputing 2011

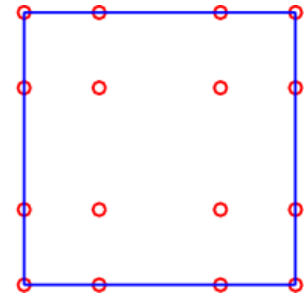
# Summary

- Using modern supercomputers to achieve high-resolution is driving our choice of grids and numerical methods
- High-resolution coupled with physics improvements leads to improved simulations
- Running climate models at climate throughput rates (~5 SYPD) at high resolution (1/8 degree) is possible with todays petascale computers and unstructured or less-structured grids

# Backup Slides

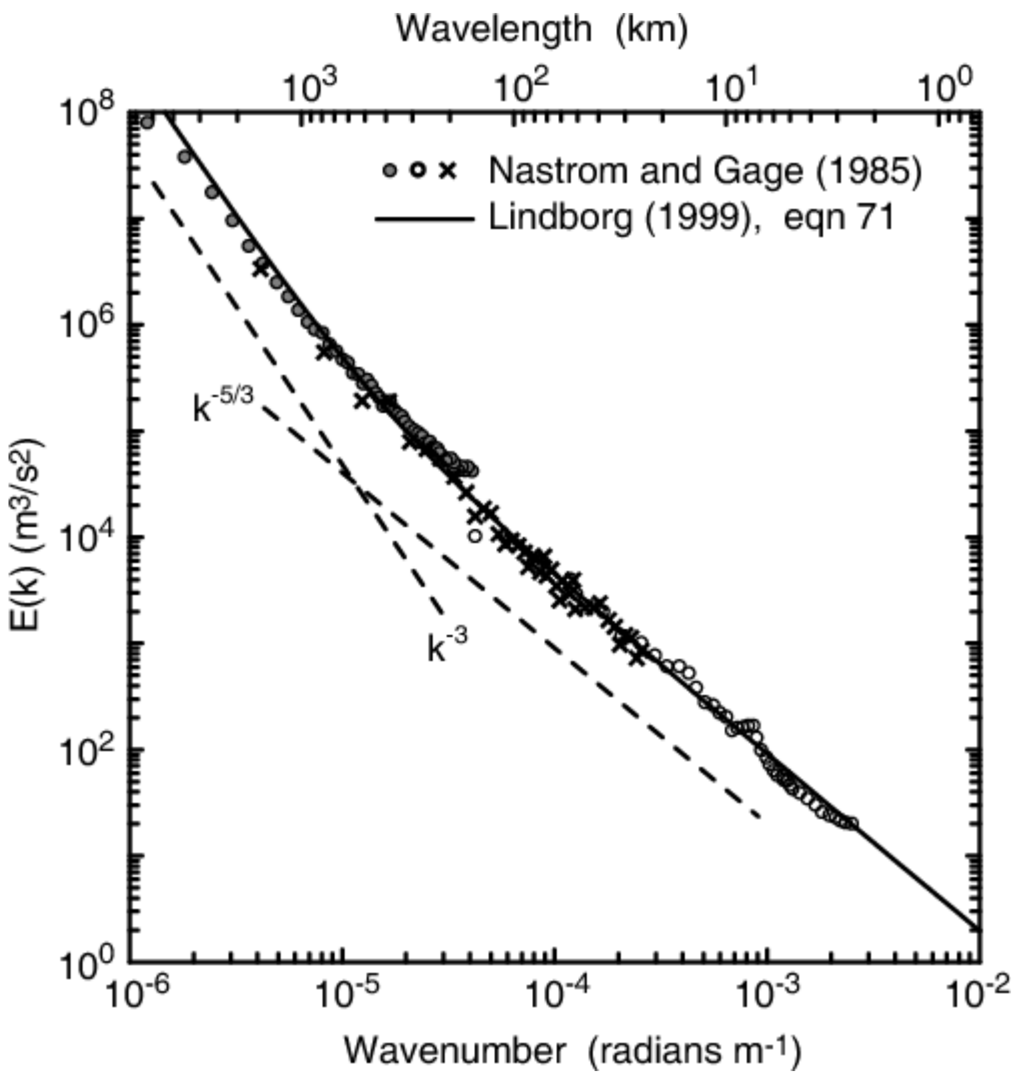
# Spectral Element Method





- Spectral Elements:  A Continuous Galerkin Finite Element Method
  - Galerkin formulation, with basis/test functions: degree *d* polynomials within each element, continuous across elements
  - Gauss-Lobatto-Legendre quadrature based inner-product (requires quadrilateral elements)
  - GLL inner product + nodal basis gives a diagonal mass matrix. (Maday & Patera 1987)

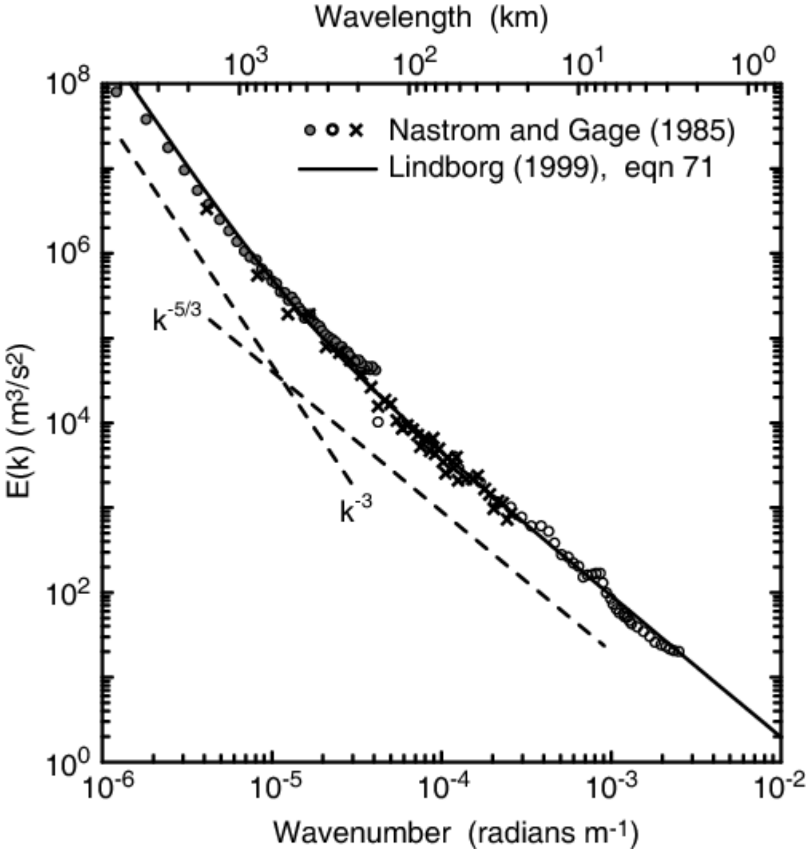
# Kinetic Energy Spectra

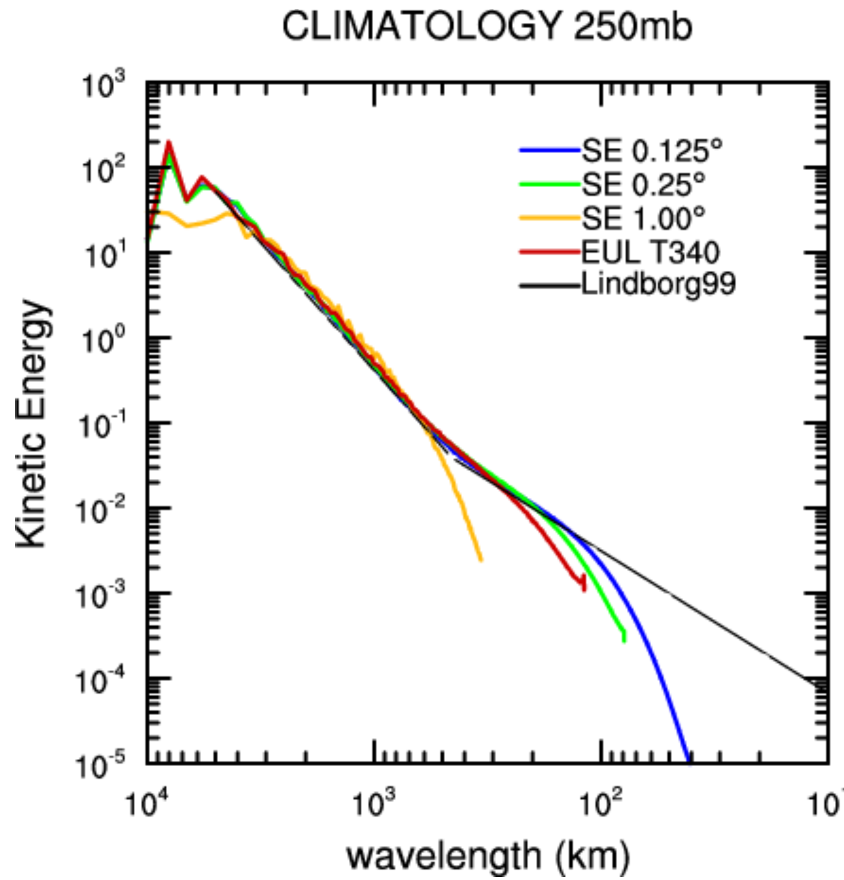



KE spectra from aircraft observations (symbols, Nastrom and Gage 1985) and functional fit (solid line, Lindborg, 1999). Figure from Skamarock, 2004.

- Nastrom-Gage transition in KE spectra
- Mesoscale shallowing:
- Transition from a -3 regime (representative of quasi-2d large scale flow) to a -5/3 regime (associated with increased variability, increased frequency of extreme events)
- Resolving the -5/3 regime considered necessary if not sufficient to simulate correct mesoscale variability
- Determine effective resolution, following Skamarock 2004

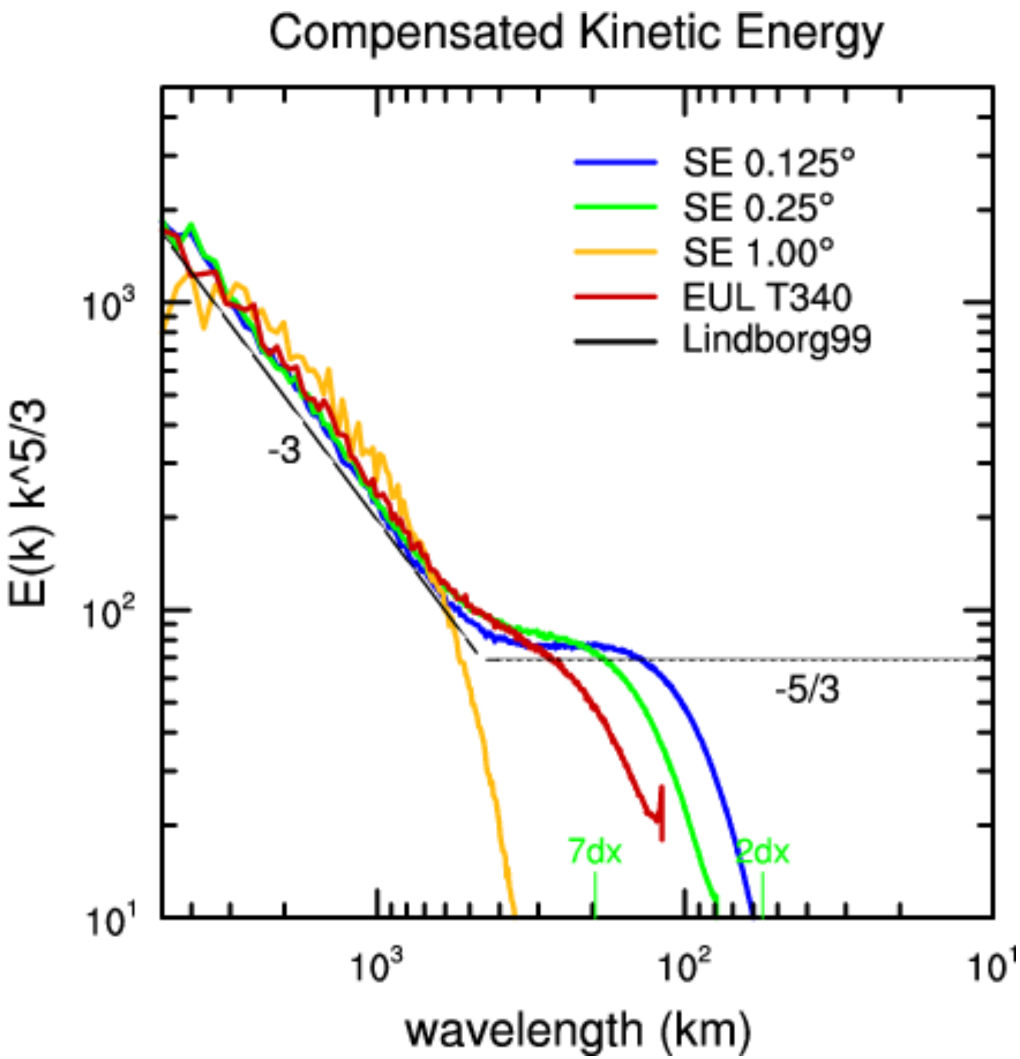
# Kinetic Energy Spectra



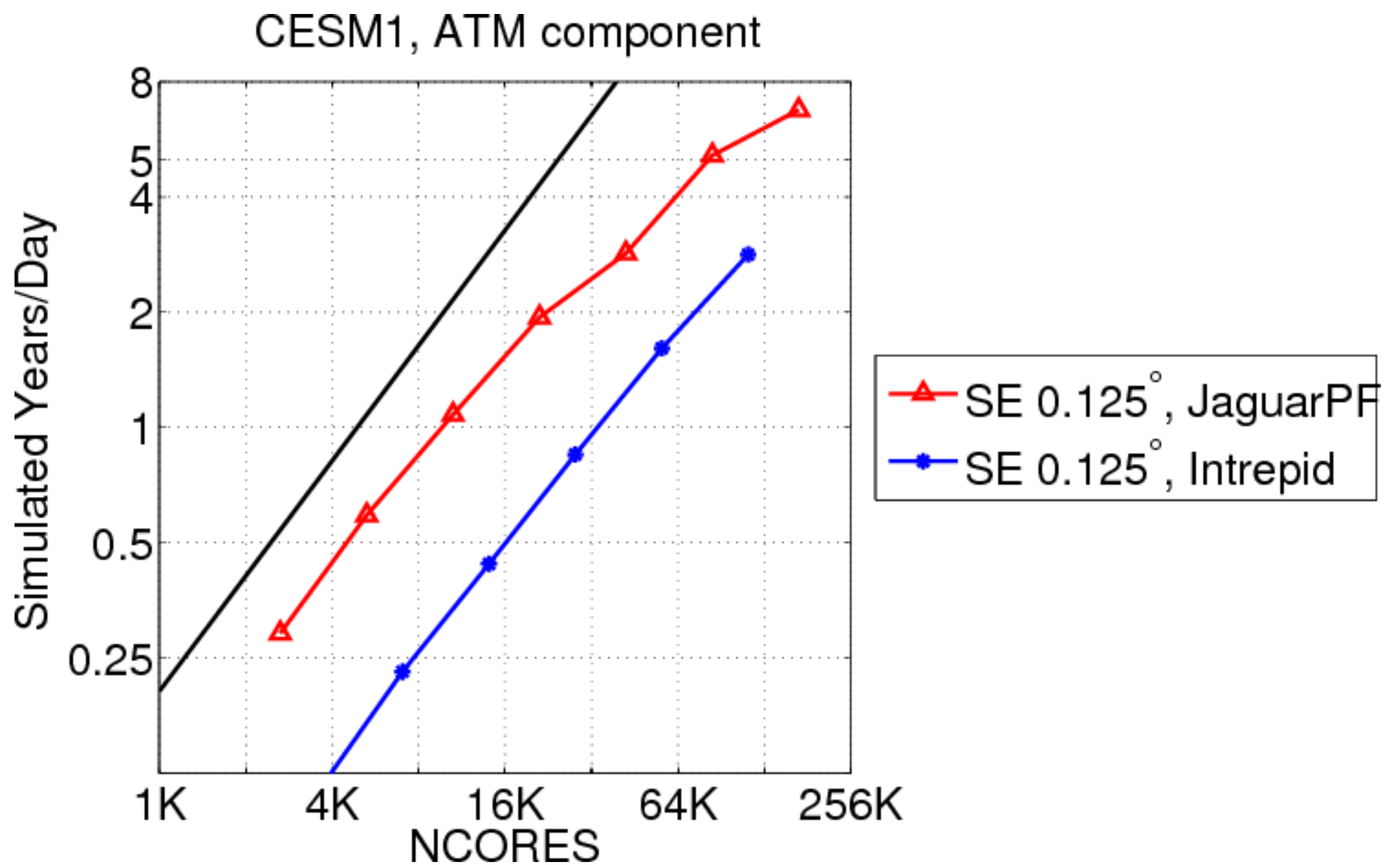

- Observations: KE transitions from a -3 scaling to -5/3, which continues to ~2km and probably even further.
- Define effective resolution as wavelength where models depart from observations

# High Resolution




- CAM minimum resolution to resolve -5/3 regime: 1/8°
- CAM 1/8° departs from observations ~100km wavelenght (50km grid scale).

# CAM4 1/8° (14km) Scalability
# IBM BG/P "Intrepid" and Cray XT6 "JaguarPF"



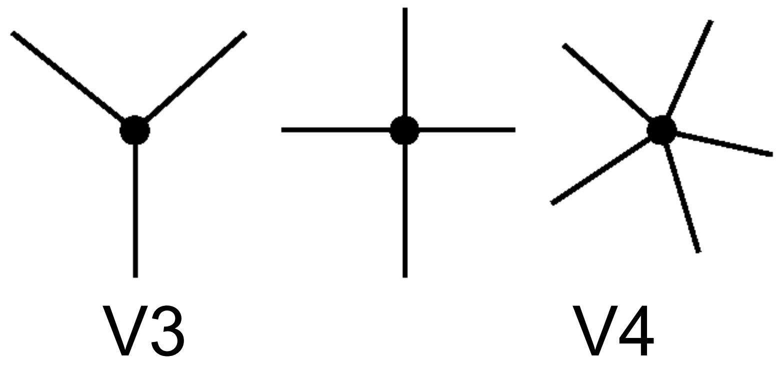

- Excellent scaling to near full machine on both LCFs:
- Intrepid (4 cores/node): Excellent scalability, peak performance at 115K cores, 3 elements per core, 2.8 SYPD.
- JaguarPF (12 cores/node): Good scalability, peak performance at 172,800 cores (2 elements per core), 6.8 SYPD.

## Euler's Formula for polyhedra: V – E + F = 2

V = number of verticies
E = number of edges
F = number of faces

Quadrilateral elements:

$$E = 4F/2$$
$$2E = \sum j \, V_j$$




Then: $V3 = 8 + V5 + 2V6 + 3V7 \,..$

most uniform solution:
$V3 = 8$, $V4$=unlimited, $V5=V6=...=0$

For non-overlapping quadrilateral grids,
The cubed sphere is the only reasonable choice!

# Conformal Cubed-Sphere Grids

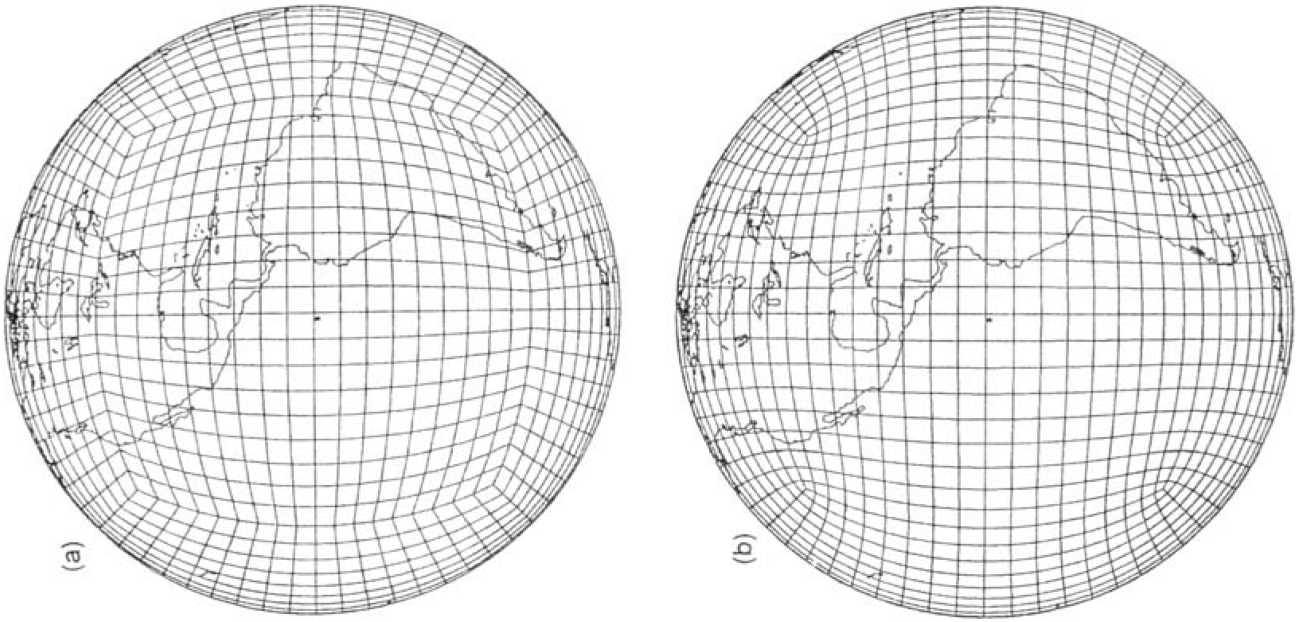



Figure 1. (a) Spherical-cubic grid using the gnomonic projection. Adjacent grid lines subtend equal angles at the centre of the sphere. (b) Spherical-cubic grid obtained by conformal transformation to map domain consisting of the square faces of the unit cube. Grid lines are equally spaced in the map coordinates.

Rancic, Purser, Mesinger QJRMS 1996
McGregor Atmos. Ocean 1996

# Conformal Cubed-Sphere Grids

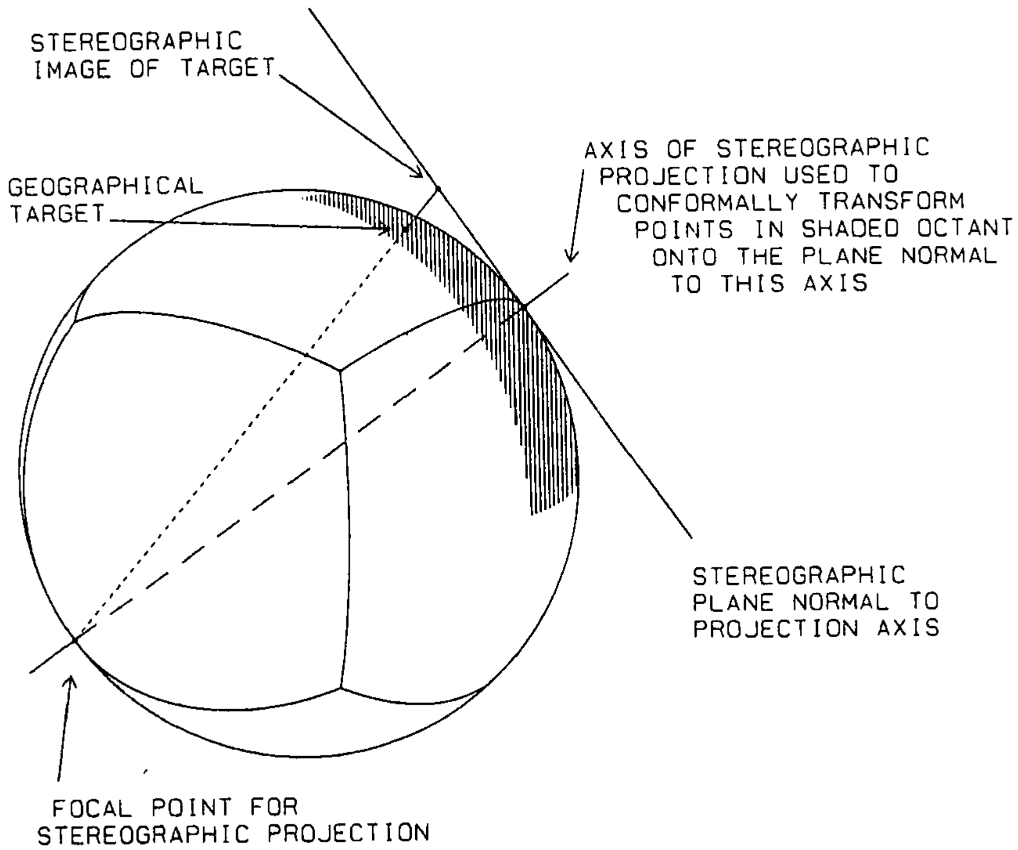



Figure A.1. Schematic depiction of the construction of the image on a plane of a target on the sphere. The image plane for any target is parallel to the sphere's surface at the point corresponding to the nearest corner of the inscribed cube. Therefore, the points on the sphere that share a common stereographic projection comprise an octant (shaded region).

Rancic, Purser, Mesinger QJRMS 1996

# Conformal Cubed-Sphere Grids

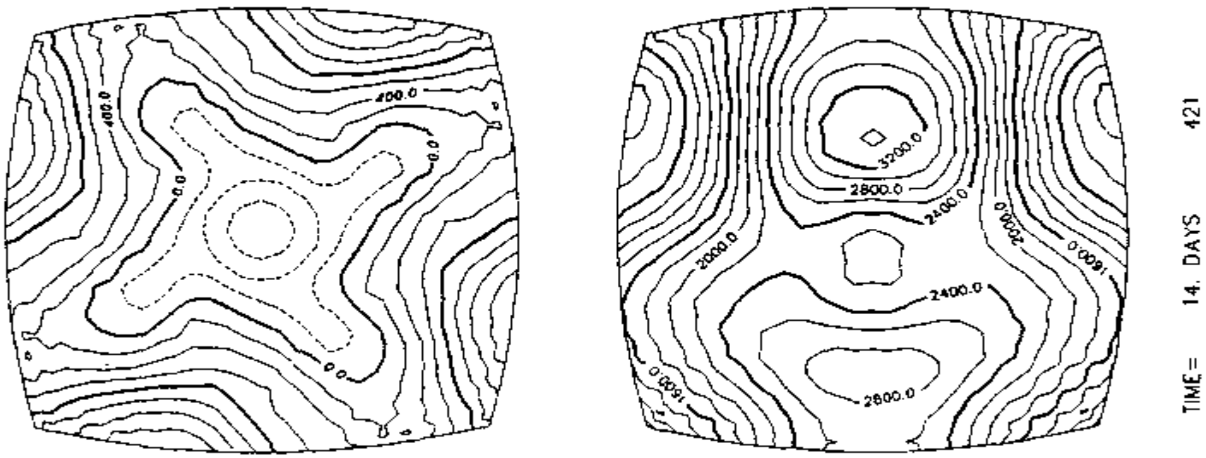
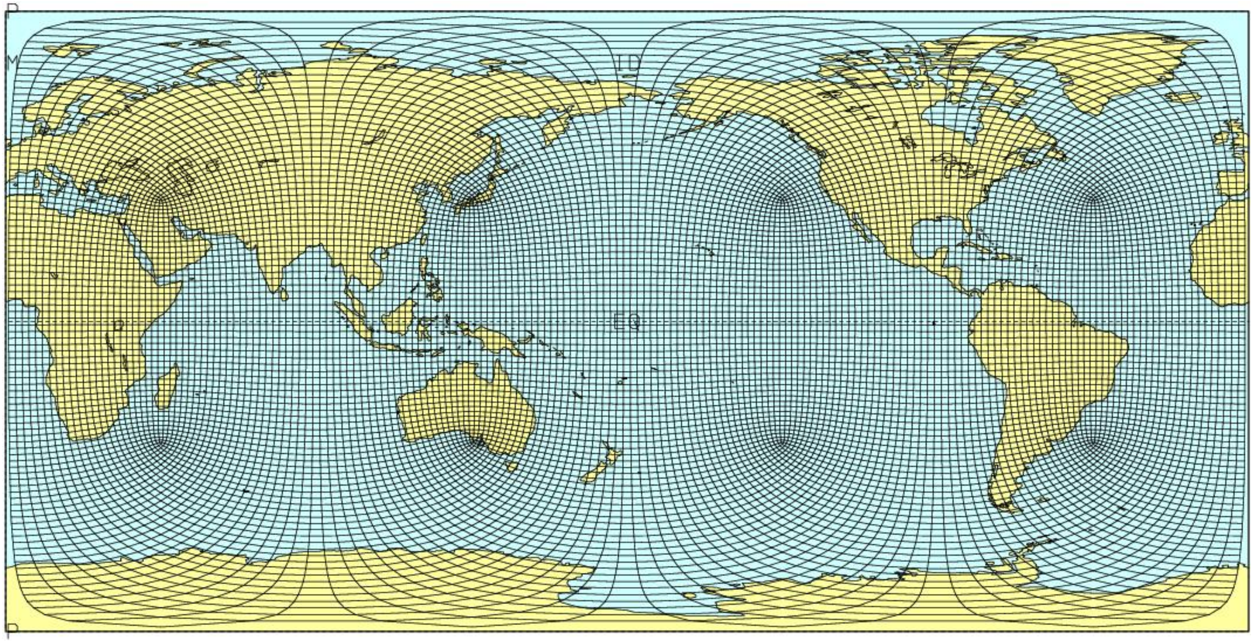



Figure 5. Numerical solution after 14 days of integration with the Rossby–Haurwitz zonal wave-4, using the E-grid, with the resolution of 421 scalar gridpoints on each of the faces of the spherical cube. The upper face (with the north pole in the middle) and one of the lateral faces are shown in the stereographic projection.

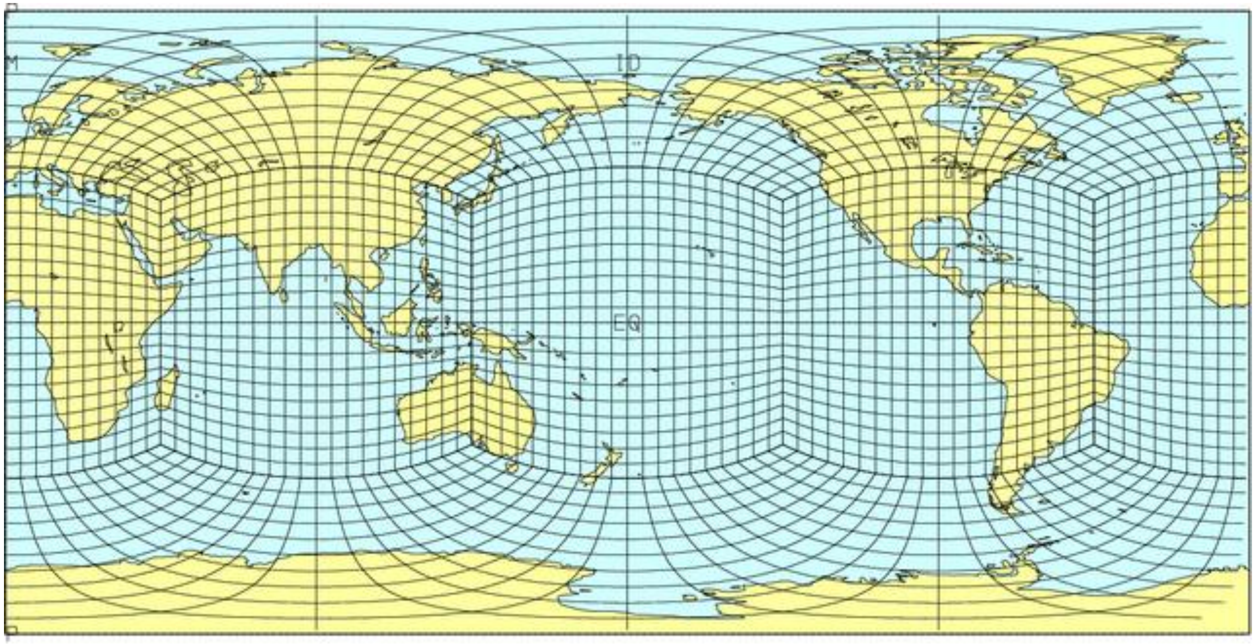
Sample cubed-sphere output from
Rancic, Purser, Mesinger QJRMS 1996

Sandia National Laboratories

# Cubed-sphere Projections




Conformal
8 “poles”




Gnomonic:
Non-orthgonal
coordinate
system

Source: J. McGregor (CSIRO) *Some features of the dynamical formulation of CCAM,*
PDEs on the sphere, 2006